Title of the Dissertation:

Modeling the climate change impacts on global coffee production

Dissertation

for the completion of the academic degree

Doctor rerum agriculturarum

(Dr. rer. agr.)

submitted to the
Faculty of Life Sciences
at Humboldt-Universität zu Berlin

by
Bunn, Christian (M.Sc.)

President of the
Humboldt-Universität zu Berlin
Prof. Dr. Jan Hendrik Olbertz

Dean of the
Faculty of Life Sciences
Prof. Dr. Richard Lucius

Advisors: 1. Prof. Dr. Dr. h.c. Dieter Kirschke
2. Prof. Dr. Hermann Lotze-Campen
3. Prof. Dr. Martin Kappas

Date of Oral Exam: 16.10.2015

Summary

To model the impacts of climate change on global coffee production in an integrated economic-biophysical modeling framework was the objective of this thesis. Coffee stands out form other crops because of its high climate sensitivity and the importance of trade for coffee markets. The vast majority of coffee is produced using either one of two species which form a single market: the heat sensitive *Coffea arabica* (Arabica) and the cold sensitive *Coffea canephora* (Robusta). Recently, evidence is increasing that climate change has begun to affect production.

Existing methodologies for climate change impacts assessments had to be improved to achieve the thesis objective. Previously, no globally coherent biophysical impact study for both crops existed even though regional studies suggested drastic impacts. To address this, machine learning classification was used to develop a global biophysical impacts model for both coffee species. Integrating these biophysical effects with demand side effects required a detailed understanding of the spatial distribution of coffee production. Because existing datasets were found to be insufficient a novel methodology was developed that built upon the machine learning classification of coffee suitability. These two steps were preconditions to include a model of the coffee sector in the spatially explicit partial equilibrium modeling framework Globiom.

On only half the area that is currently available for coffee production by 2050 2.5-times as much coffee will have to be produced to meet future demand. Reduced yields and increased prices were shown to reduce the coffee market by more than 5million tons per year, equivalent to the size of the baseyear 2000 market volume. Coffee production will migrate to higher elevations where area is available for agricultural production. Production will remain within current latitudinal ranges but major producers like Brazil and Vietnam will struggle to remain competitive with relatively less affected countries in East Africa. Substantial uncertainty about the impacts on local scale prevails and impedes the development of unambiguous adaptation strategies. Thus, there will be coffee on the table in 2050, but it will be of lower quality, will cost more and it will still be in the focus of sustainable enterprises because its production continues to be shaped by poverty risk and environmental problems.

Zusammenfassung

Die Untersuchung der Auswirkungen des Klimawandels auf die globale Kaffeeproduktion in einem integriertem ökonomisch-biophysischem Modell war das Ziel dieser Arbeit. Kaffee unterscheidet sich von anderen Kulturarten durch die starke Klimaabhängigkeit und der Bedeutung des Handels für Kaffeemärkte. Der vorwiegende Teil der globalen Kaffeeproduktion stammt von zwei Arten: dem hitzeempfindlichen *Coffea arabica* (Arabica) Strauch und vom frostempfindlichen *Coffea canephora* (Robusta). Eine zunehmende Zahl Studien zeigt, dass der Klimawandel bereits heute die Produktion mindert.

Existierende methodische Ansätze zur Schätzung der Folgen des Klimawandels wurden im Rahmen dieser Arbeit erweitert. Obwohl regionale Studien drastische Auswirkungen aufgezeigt hatten, war vorher noch keine global einheitliche Studie zu den biophysischen Effekten für beide Kaffeearten durchgeführt worden. Maschinenlernklassifizierung wurde hier genutzt um ein Modell der globalen Klima-Kaffee-Wechselwirkungen zu entwickeln. Zur Integration der modellierten Klimafolgen mit ökonomischen Faktoren war ein detailliertes Wissen über die räumliche Verteilung der Kaffeeproduktion notwendig. Da existierende Datensätze unzureichend waren, wurde ein neuer methodischer Ansatz auf der Grundlage der maschinenlern-basierten Anbaueignungsklassifizierung entwickelt. Diese beiden Schritte waren Voraussetzung für die Inklusion eines Modells des Kaffeesektors in dem räumlich expliziten partiellen Gleichgewichtsmodell Globiom.

Auf der Hälfte der heute für den Anbau geeigneten Fläche muss bis 2050 2,5-mal so viel Kaffee produziert werden um die zukünftige Nachfrage zu sättigen. Niedrigere Ernten und höhere Preise werden das Volumen des Kaffeemarktes um über 5 Mio. Tonnen pro Jahr reduzieren. Dieser Verlust entspricht dem Marktvolumen im Modellbasisjahr 2000. Kaffeeproduktion wird zukünftig in höheren Lagen angebaut werden müssen, sofern dort landwirtschaftliche Fläche zur Verfügung steht. Die Produktion wird größtenteils innerhalb der gegenwärtigen Breitengrade bleiben, aber wichtige Produzenten, wie Brasilien und Vietnam werden Probleme haben wettbewerbsfähig zu bleiben mit weniger betroffenen Ländern in Ost-Afrika. Modellunsicherheit auf lokaler Ebene erschwert jedoch die Entwicklung eindeutiger Anpassungsempfehlungen. Es wird also auch in Zukunft Kaffee geben, aber dieser Kaffee wird von geringerer Qualität sein, mehr kosten, und er wird immer noch das Lieblingsobjekt von Nachhaltigkeitsunternehmern sein, da die Produktion auch weiterhin von Armutsrisiko und ökologischen Problemen geprägt sein wird.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

AEZ	Agro Ecological Zoning
AR4	4th Assessment Report of the IPCC
AR5	5th Assessment Report of the IPCC
AUC	Area Under the receiver characteristic Curve
CaNaSTA	Crop Niche Selection for Tropical Agriculture
CATIE	Centro Agronómico Tropical de Investigación y Enseñanza
CGE	Computable Equilibrium Model
CIAT	International Center for Tropical Agriculture
CRU	Climate Research Unit
FAO	Food and Agriculture Organization of the United Nations
GAEZ	Global Agro-Ecological Zoning
GDP	Gross Domestic Product
HRU	Homogenous Response Unit
IBGE	Instituto Brasileiro de Geografia e Estatística
ICA	International Coffee Agreement
ICO	International Coffee Organization
IFPRI	International Food Policy Research Institute
IIASA	International Institute of Applied Systems Analysis
INEC	Instituto Nacional de Estadística y Censos (Costa Rica)
IPCC	United Nations Intergovernmental Panel on Climate Change
NoCC	No Climate Change
NPV	Net Present Value
RCAM	Central America Region
RCP	Representative Concentration Pathways
SimU	Simulation Unit
SRES	Special Report on Emission Scenarios
SVM	Support Vector Machines
USDA	United States Department of Agriculture

Acknowledgements

This work would not have been possible without the support of many people. I am especially grateful for the trust that I have experienced and the freedom that I was given to pursue my ideas.

I am highly indebted with my colleagues at the International Center for Tropical Agriculture (CIAT) in Cali, Colombia, the International Institute of Applied Systems Analysis (IIASA) in Laxenburg, Austria, and my thesis supervisor Prof. Dr. Dr. hc. Dieter Kirschke, Humboldt-Universität zu Berlin, Germany. While working together over such remote distances I was oftentimes dependent on the trust and continued support of my colleagues when we did not see each other personally for several months or even years. Prof. Kirschke kept supporting my repeated adjustments to my thesis plans, even when these changes were mostly based in my personal curiosity. To me, such trust and freedom cannot be taken for granted and I would like to express my gratefulness for it.

From the beginning Dr. Peter Läderach (CIAT) believed in me and my ideas. Peter inspired me with his work and provided me with the means to carry it out. When I read one of his blog posts about five years ago I knew what I wanted to do for my doctoral research. I am still thankful that he gave me the opportunity to work in his team when I first applied to CIAT. During the entire thesis work he provided me with the intellectual and financial support necessary to conduct the research presented here. I sincerely hope that in the coming years we will be able continue our work on our common passion.

This thesis would have been much less exciting had not Dr. Peter Läderach and Dr. Michael Obersteiner (IIASA) shared a taxi from a conference to the airport. Had they not discussed my ideas and decided to support my work, I would not have achieved the results here. It was because of this chance encounter that I was given the opportunity to work with the Globiom team. Working with them was both challenging and highly rewarding because of their great ideas and their spirit. At IIASA Dr. Aline Mosnier helped me with great patience during my first steps with the Globiom code. Her help was crucial to achieve my thesis objectives and I look forward our continued collaboration.

The Data and Policy Analysis area (DAPA) people at CIAT keep amazing me. The DAPA leader Dr. Andy Jarvis with his personality and example is truly inspiring and I am incredibly grateful that I could experience such a team. The scientific freedom, innovative spirit, cheerfulness, friendship and positive mindset at DAPA are unmatched. This experience will shape me as a person and researcher. It inspired me to believe in my ideas and to carry on even during the more frustrating times of this thesis work. A special mention goes to the Species Distribution Modeling group in DAPA for the fruitful discussions that resulted in great improvements of this thesis.

During the course of this thesis I was never alone. Other doctoral students were always there to help, be it by revising a manuscript, their spirit or just by having nerdy conversations over beer or coffee. These people might not know that without them this would have been much harder, so I would like to thank them here: Anton Eitzinger, Theresa Liebig, Eric Rahn, Joann de Zegher, Dr. Julia Schmidt, Nora Castaneda, Colin Khoury, Wytse Vellema.

Last, but really most important, I owe my family and friends. My parents always let me follow my curiosity and supported me with whatever idea I could come up with like going to Japan, Colombia or changing my degree goals from molecular plant physiology to global climate change. Not all friendships would have survived my choices. I am therefore grateful for all the late night beers and random discussions with my friends Jonas and Henning that probably cost me some brain cells, but also inspired new ideas.

Family and friends are also my wife Beatriz Vanessa and Sara Melissa. Beatriz was my secret thesis supervisor and Sara a constant source of energy. These two provided me with all of the above: inspiring ideas, moral support, friendship, honest feedback, and a lot of patience. Yet, this thesis would still not be completed if my son would not be due in a few weeks so that I really have to come to an end.

This research was conducted under the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Part of this thesis was funded through a "Stipendium Klimafolgenforschung" of the Stiftung Humboldt-Universität. Additional funding came from World Coffee Research as part of the project "Identifying Long Term Variety Trial Locations, Provide Climate Information to Support World Coffee Research Variety Trials and Support on Trial Data Analysis."

1 Introduction

There are two things coffee producers all over the world are worried about: The weather and the coffee price. Climate change has begun to affect coffee producers, while demand for coffee is growing. Even though nearly all coffee is traded, to date the effects of these developments have not been researched in a global economic context. In this thesis an integrated modeling framework will be developed to examine how a progressing change in climate conditions will affect coffee production while also accounting for an increased demand for coffee in a globalized world.

In urban service societies coffee has become a cultural beverage that is deeply rooted in today's zeitgeist. Introduced as an exotic treat for the rich, coffee passed through a phase of being a mass consumption commodity in more industrial times. Over the past decade or so, it is increasingly becoming a lifestyle product. It is consumed in a large variety of forms that are often culturally tied, allowing the expression of personal preferences. Additionally, coffee production has become a playing field for new approaches to sustainable consumption. Little caters to the urban dwellers desire for lasting values like images of Andean smallholders carefully picking ripe coffee beans in a tropical agro-forestry system.

Thus, coffee can be a lot more than just a bitter boost in the morning; it is a link between rich consumers and poor producers. The safety nets of first world societies meet a production that is characterized by high economic and natural risks. A good harvest does not necessarily result in good returns as prices may be low. Such was the case during the decade around the millennium known as the "coffee crisis". Worse, for many producers a bad harvest may mean a hungry year. The reasons for such failure often lie beyond the farmer's reach. E.g. there may be no cure for fungal diseases like coffee rust that destroy landscapes of plantations; or events like the "El Nino" fluctuation cause large variations of rainfall and heat between seasons. Such climate risk is generally considered the largest source of uncertainty in coffee production. Now, climate change has started to affect crop production globally.

Even early accounts on the production of coffee mention the need of suitable climatic conditions (e.g. in Ward 1911). The specific climatic requirements of the coffee crop have made the topic of climatic suitability a constant in coffee cultivation. The equation for climate change impacts is simple at first glance: many sources describe the climatic requirements of coffee with reduced, very general statements such as "18°C to 22°C annual mean temperature". Climate research projects global warming of 2°C annual mean temperature until mid-century. Naively, stakeholders in the coffee industry suggested that in this case the coffee area that previously had annual mean temperatures above 20°C would be replaced by novel area with annual mean temperatures that previously were around 16°C. Unfortunately, neither climatic change, nor the climatic requirements of the coffee crop can be summarized that easily. The projected climatic changes are specific for each region and yields are compromised by e.g. heat waves rather than annual temperature means.

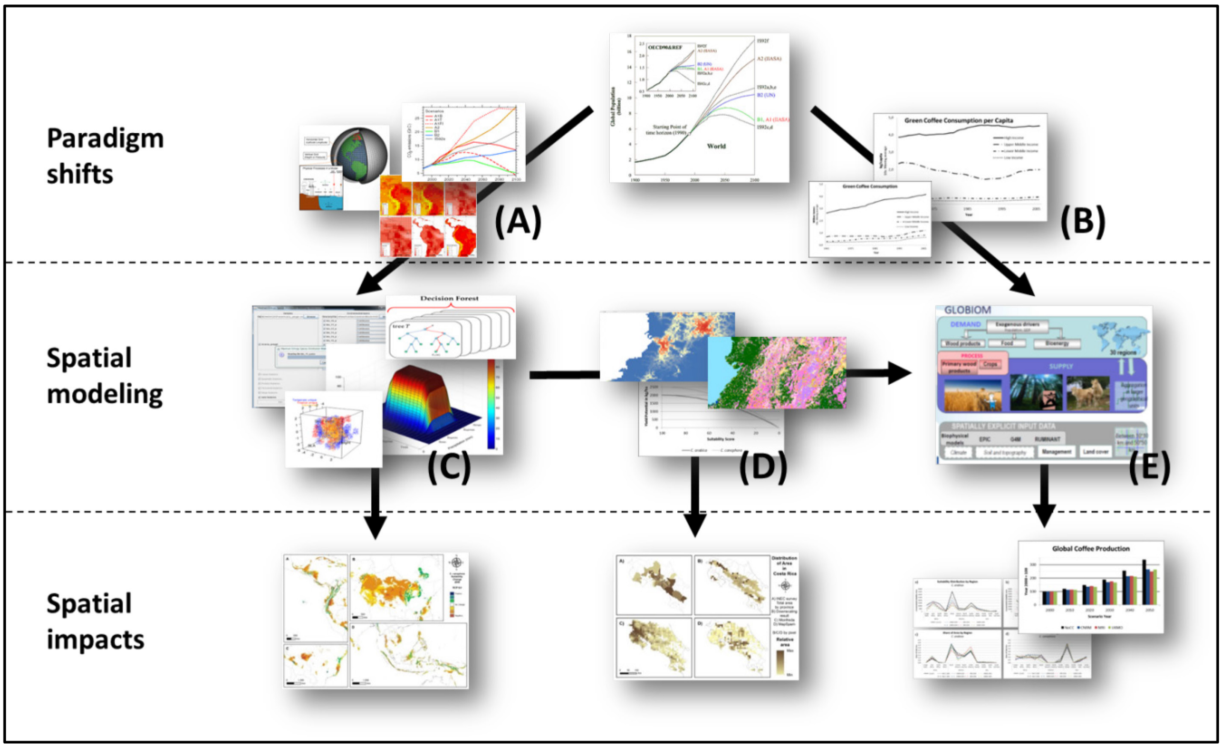
An interesting aspect of coffee is that it is largely traded, little is consumed at origin. The vast majority of coffee is produced using either one of two species: *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta). The products of the two species form a single market on which they act as substitutes. But while Arabica is of higher quality and commands a price premium, Robusta is more productive. Coffee smallholders are therefore interconnected through global markets (Eakin, Winkels and Sendzimir 2009). Unlike staples coffee production is not relevant from a global food security perspective. Regionally however, coffee production has shaped entire landscapes. In these regions large shares of the rural workforces depend on labor in coffee, and smallholder families rely entirely on coffee income. In extreme cases revenue from coffee exports is an important source of foreign income for entire countries.

The global coffee market has been characterized by steadily increasing demand over the past decades. This increase came from two sources: in saturated developed markets in Europe and USA the increase came from population increases with constant per capita demand. Since the end of the quota regime of the International Coffee Agreement (ICA) world market prices have fallen and per capita consumption increased also in emerging middle income countries. The additional demand was met by increased per hectare productivity in producing countries, while overall acreage remained constant. Over the past decades African regions were on the decline and coffee area was reduced, likely because of low productivity. On the other hand highly productive producers in Asia replaced the African area.

The paradigms of the coffee sector have been projected to shift fundamentally in coming decades as a consequence of climate change and its driving forces. Emissions that result in anthropogenic climate change are mainly caused by population growth and economic development. Changes in these paradigms can be expected to have strong repercussions in global coffee production, because the coffee sector is characterized by high regional concentration of production, determined by the suitability of the climate for coffee production, and driven by population and income increases demand is already increasing globally.

The main objective of this thesis is the ex-ante evaluation of the impacts of the projected changes in the fundamental supply and demand paradigms of coffee production in a systematic sequential modeling framework (Figure 1). In sequential modeling approaches the results of one model feed into another model in a hierarchical modeling framework. This framework will also structure this thesis.

Figure 1. Modeling framework for climate change impacts on coffee



A) Global climate models; B) population and economic growth change demand; C) machine learning classification for biophysical impact surfaces; D) generation of spatial production data using suitability maps; E) integrated economic-biophysical impact scenarios using Globiom (own representation with graphics from Solomon et al. 2007; Jarvis and Ramirez 2010; Havlík et al. 2011).

The projections of global climate models (GCM) that are driven by scenarios of anthropogenic greenhouse gas (GHG) emissions (Figure 1.A) are the topic of chapter 2.1. A brief review of trends in coffee consumption and production in chapter 2.3 provides the foundation for future demand scenarios (Figure 1.B). The remainder of chapter 2 highlights two gaps in the literature: Previous work that related progressive climate change to coffee production was confined to regional scale (chapter 2.2). Research using integrated biophysical-economic modeling frameworks previously did not consider coffee even though the larger share of the global coffee production is traded. Economic effects could alter the outcomes of climate change substantially. Therefore, a coffee model will be included in the Globiom framework (chapter 2.4). To achieve the main objective of ex-ante evaluating the impacts of climate change related paradigm shifts three secondary objectives had to be addressed.

The first contribution of this thesis is the demonstration of the use of machine learning algorithms to estimate the global impacts on the climatic suitability for coffee production (Figure 1.C). Crop models in climate change impact assessments attempt to quantitatively represent the climate-plant interaction. The projected impacts of climate change on coffee have been discussed in numerous studies before the start of this thesis. However, none of the pre-existing studies had assessed the impact on global scale in a coherent way for both coffee crops (Chapter 2.2 - Impact models for coffee). Therefore, chapter 3 (Climate change profile of global suitability for Arabica and Robusta coffee) will develop a global model for both coffee crops. As a result our understanding of the impacts of climate change on the global suitability for coffee production has been advanced.

The second contribution is the disaggregation of national coffee production statistics to probable locations of production using a novel approach (Figure 1.D). Climate change impacts on suitability from chapter 3 did not equate impacts on coffee production. The biophysical impacts at any given location have to be related to the area under coffee production at that location to be able to focus on changes in climatic suitability relevant to the sector. Existing datasets of spatially explicit production data were found to be too unspecific for the purpose here (Eriyagama, Chemin and Alankara 2014). Therefore the suitability data from chapter 3 was used to develop maps of harvested area statistics (Chapter 4 - Where on earth is coffee grown? Spatial disaggregation of harvested area statistics using suitability data). These maps reflect with improved accuracy the physical distribution of coffee production at a spatial resolution that make them useful for coffee specific studies.

The third contribution then directly leads up to the overarching objective. The final objective of this thesis was to integrate the supply side changes caused by climate change and the demand side changes from a changing global demography in the partial equilibrium framework Globiom (Figure 1.D). Precondition for the integration of coffee data in Globiom was the estimation of global climate change impacts on yields in chapter 3 and the availability of accurate harvested area data from chapter 4. Demand scenarios were based on the demographic and economic story line of the climate change scenario for a coherent modeling framework. This allowed the development of impact scenarios of changes in the paradigms of coffee production in future periods (Chapter 5 - Integrated biophysical-economic assessment of the climate change impacts on global coffee production).

2 Changing paradigms for coffee production: What the past tells us about the future

This thesis will integrate two aspects of coffee production in a sequential modeling framework: The supply side shifts in production caused by climate change and the demand side shifts caused by the corresponding changes in global population and incomes. In this chapter the hypothesis that coming changes will change the paradigms of the coffee sector will be further elaborated based on previous research. The background to the guiding objective is provided by the projections of global circulation models. First, some aspects of modeling climatic changes are reviewed. The importance of climate for coffee production is demonstrated and previous studies on the impacts of climate change on coffee are reviewed. Then, a historic perspective on coffee production is taken. Past developments of production and consumption are discussed as this information will later be used to develop future demand scenarios. Last, the Globiom integrated modeling framework is introduced to show how changes in the paradigms may be integrated.

The first section on historical and future climate will not focus as much on the projected climatic changes themselves but on the sequence of data used in the framework used here. Both ends of this modeling sequence are relevant for this thesis. The final output of the process is high resolution climate data that is spatially disaggregated to a scale relevant to agricultural production. The biophysical impact model will use climate data of current conditions on the same scale as a baseline for comparison. Therefore also this data will be discussed here as it is a necessary step to model supply side impacts on coffee production. However, the starting point are the emission scenarios that are used to drive global circulation models, the representations of the global climate system. The emission scenarios themselves are based on assumptions about future population growth and economic pathways. They therefore provide the background for the demand side changes in the coffee sector.

Both species that are predominantly used for coffee production depend on the climate for high productivity. Arabica coffee is more susceptible to higher temperatures than Robusta. But the distribution of Robusta is limited by its low tolerance for low temperatures. With anthropogenic climate change increasingly being more than an academic research question, but a reality with undeniable effects, the question how this will affect coffee producers is self-evident. Therefore, the projected impacts of climate change on coffee have been discussed in numerous studies before the start of this thesis. As will be demonstrated, three shortcomings can be identified: None of the pre-existing studies had assessed the impact on global scale in a coherent way for both coffee crops. And, the majority of these studies looked exclusively at changes in the climatic conditions for coffee production, with only few exceptions linking these changes to economic indicators. Furthermore, none of the previously existing studies had incorporated demand side changes into the analysis even when regarding multi-decadal time horizons.

The section on coffee consumption and production will provide support for the hypothesis that the projected changes in climate and population will impact the coffee sector. It will be shown that demand changes in the past decades have been driven mostly by population growth in saturated markets. Over the past decades demand in emerging economics has picked up substantially which can be associated with income growth. Producers met this increased demand mostly by substantially increased productivity. African origins that could not compete were left behind.

Climate change has started to affect crop production globally. Both coffee species are highly climate dependent, but with different characteristics. Through global coffee markets the two crops are interconnected. Market fluctuations are often drastic, with direct consequences for smallholder farmers. An analysis of the climate change impacts on global coffee production would be incomplete without taking into consideration market effects. It was therefore proposed to model these effects in an integrated fashion using a partial equilibrium model of the agricultural sector. In comparison with other models of the same kind Globiom has several advantages: Globiom is global, spatially explicit with high disaggregation, non-monetary flows are included, and in addition to crop production also the livestock sector is modeled and a forestry model is attached. The last section of this chapter will introduce the Globiom modeling framework and discuss which steps are necessary to include coffee in this framework.

2.1 Historical and future climate

Understanding the climate of the future requires an understanding of the historic climate. Not only are GCMs calibrated to reproduce historic climate development, also crop models are trained and calibrated against such data. This section therefore starts with a review of spatial data of the baseline climate before briefly discussing the process of projecting climatic changes using GCMs.

2.1.1 Current climate surfaces

Spatially explicit climate change impact modeling usually includes some kind of comparison of future conditions with “today” or “current” conditions. Therefore just as important as knowledge about future developments is solid knowledge on the current conditions. Such data is used as reference baseline data. It is meant to reflect a climate unaffected by anthropogenic influences. It serves both for comparisons of changes and as a major data source for model implementation.

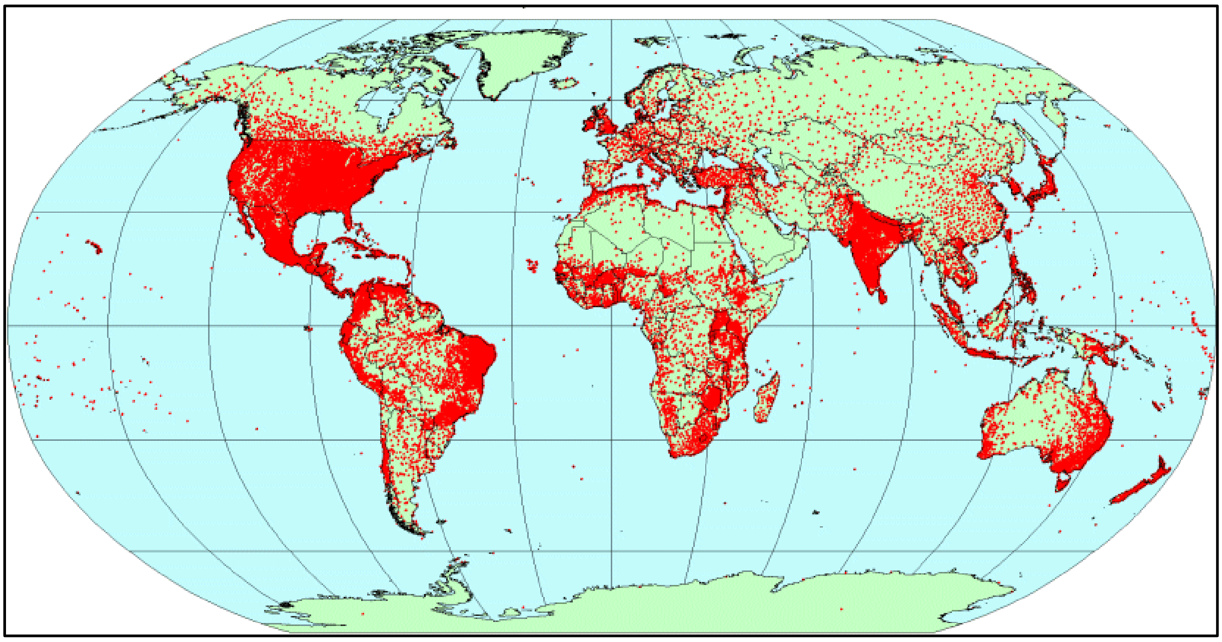
The most common approach to generate surfaces of climate data is the interpolation of historic data. Weather information from regional or global datasets of local weather stations is spatialized. Gaps between weather stations are filled by different interpolation methods, such as inverse distance angular weighted interpolation or similar approaches. Often a correction for altitude is applied to reflect topographical effects.

Regional studies often assemble a tailored dataset to have full control over the climate information used, the interpolation method and quality requirements. On global scale however such an effort is highly limited by computational requirements. Many coffee impact studies have therefore resorted to publicly available datasets of climate surfaces, namely WorldClim (Hijmans et al. 2005) and Climate Research Unit (CRU) data (Mitchell et al. 2004). Here, we briefly introduce this data.

2.1.1.1 WorldClim

One of the most commonly used data bases for the current climate (1950–2000) is the WorldClim global climate data set (Hijmans et al. 2005). The interpolated climate surfaces for global land areas were generated from a comprehensive set of climate data sources for the globe. Data is interpolated using the Anusplin thin plate smoothing spline algorithm using latitude, longitude and elevation as independent variables. Resulting are surfaces of monthly precipitation and mean, minimum, and maximum temperature on 5”, 2.5”, and 0.5 arcmin scale. The database is based on precipitation records from 47 554 locations, mean temperature data from 24 542 locations, and minimum and maximum temperatures from 14 835 locations with uneven geographic distribution. Especially regions with low population density are underrepresented in the station data. Even though precipitation data coverage is rather good, especially typical *C. canephora* production regions are not necessarily well represented, e.g. in Brazil’s Rondonia region or the Congo basin (Figure 2).

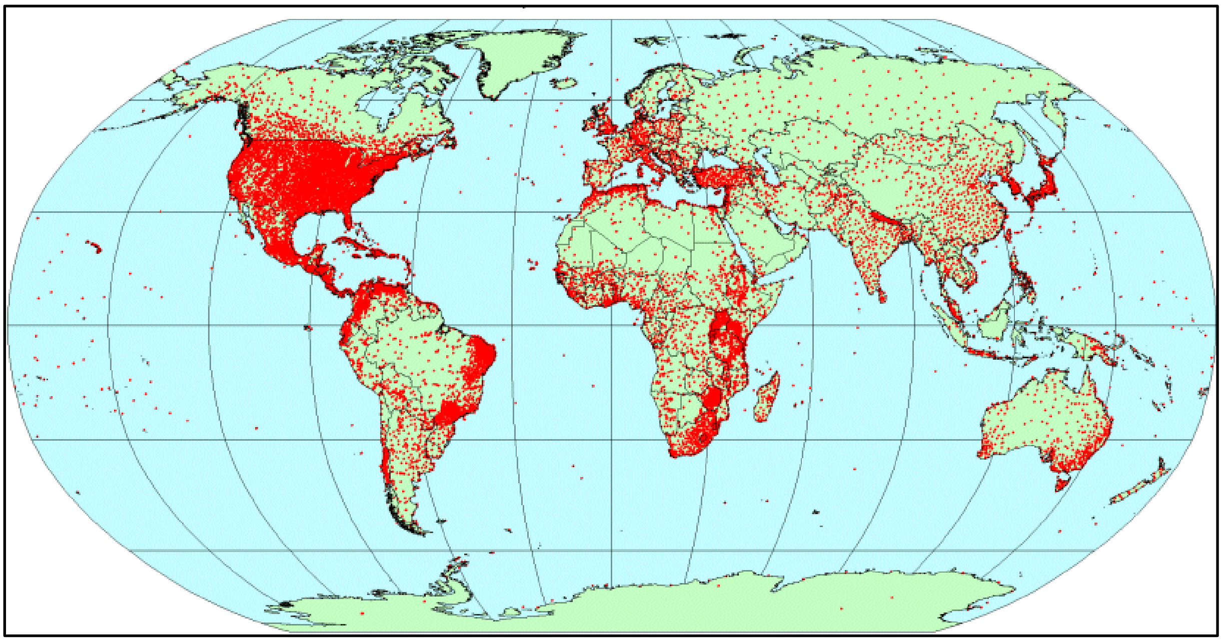
Figure 2. Locations of climate stations with precipitation data



(Taken from Hijmans et al. 2005)

In addition to the latter regions mean temperature data is also sparse in Indonesian islands and Western Africa (Figure 3).

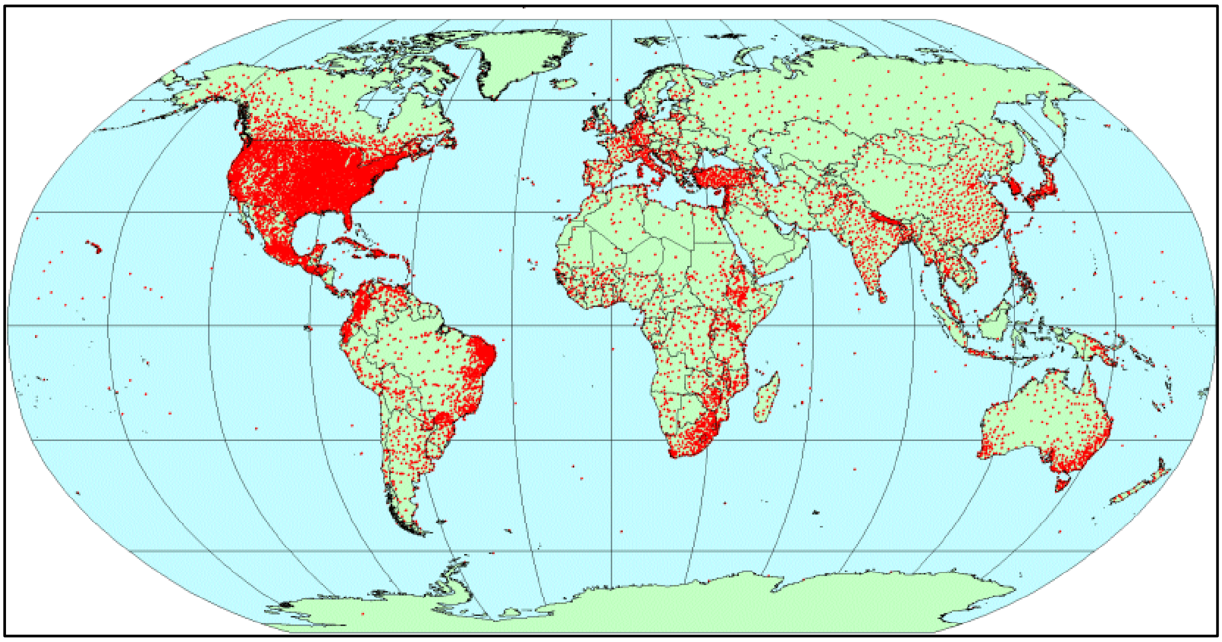
Figure 3. Locations of climate stations with mean temperature data



(Taken from Hijmans et al 2005).

Data for temperature range is even sparser. Single stations are used to interpolate on entire regions, even such climatically heterogeneous zones as the Andes (Figure 4). A notable limitation from a coffee perspective is also the sparsity of data in West and Central Africa.

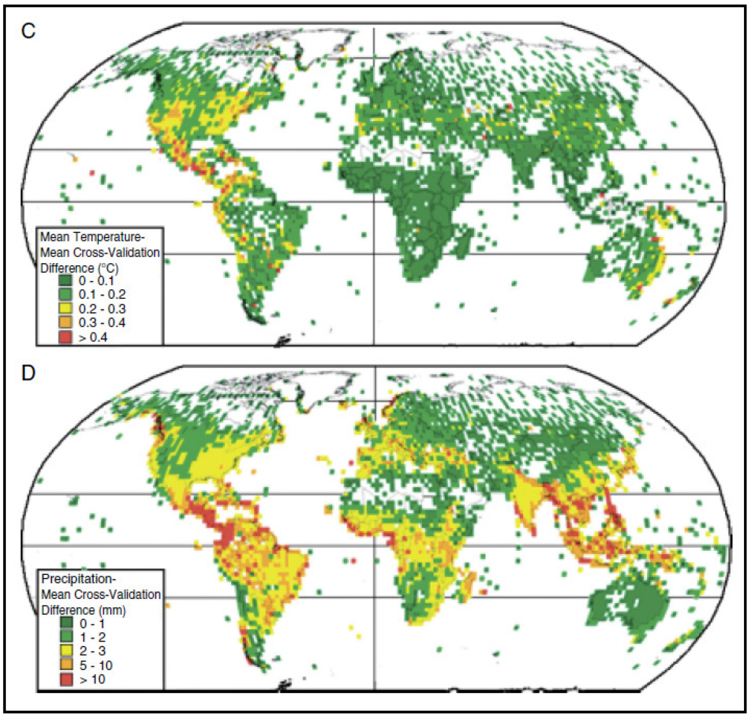
Figure 4. Locations of climate stations with temperature range data



(Taken from Hijmans et al. 2005).

From a coffee perspective the data could therefore be better. In general, uncertainties are criticized to be rather high in high altitudes for precipitation data. The important coffee origin Central America shows some of the highest uncertainties for temperature data. Also, mean deviations for precipitation in tropical regions are often higher than 10mm/month (Figure 5).

Figure 5. Uncertainty in the climate surfaces



The figure shows mean cross-validation deviations when partitioning data in test and training sets for (C) temperature and (D) precipitation averaged over 12 months by 2 degree grid cell (taken from Figure 3 in Hijmans et al. (2005).

2.1.1.2 BioClim

In addition to the monthly precipitation and temperature data the WorldClim dataset provides interpolated climate layers for 19 bioclimatic variables based on historical data (Table 1). These variables represent patterns found in the monthly weather station data, e.g. annual temperature and precipitation extremes, seasonality and means. This approach was originally perceived by (Nix 1986). Additional variables can be generated if they require no more than monthly mean temperatures, mean maximum temperature, mean minimum temperature, or mean monthly precipitation. E.g. (Läderach et al. 2013) propose a method to derive evapotranspiration bioclimatic variables that proved useful in their cocoa model. However, the most commonly used variables are the 19 original variables.

Table 1. List of bioclimatic variables available from WorldClim

Type	Bioclimatic variable	Description	Unit
Temperature	BIO 1	Annual Mean Temperature	°C
	BIO 2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	°C
	BIO 3	Isothermality (BIO2/BIO7) (*100)	-
	BIO 4	Temperature Seasonality (standard deviation *100)	°C
	BIO 5	Max Temperature of Warmest Month	°C
	BIO 6	Min Temperature of Coldest Month	°C
	BIO 7	Temperature Annual Range (BIO5-BIO6)	°C
	BIO 8	Mean Temperature of Wettest Quarter	°C
	BIO 9	Mean Temperature of Driest Quarter	°C
	BIO 10	Mean Temperature of Warmest Quarter	°C
	BIO 11	Mean Temperature of Coldest Quarter	°C
Precipitation	BIO 12	Annual Precipitation	mm
	BIO 13	Precipitation of Wettest Month	mm
	BIO 14	Precipitation of Driest Month	mm
	BIO 15	Precipitation Seasonality (Coefficient of Variation)	-
	BIO 16	Precipitation of Wettest Quarter	mm
	BIO 17	Precipitation of Driest Quarter	mm
	BIO 18	Precipitation of Warmest Quarter	mm
	BIO 19	Precipitation of Coldest Quarter	mm

(Hijmans et al. 2005)

2.1.1.3 Climate Research Unit data

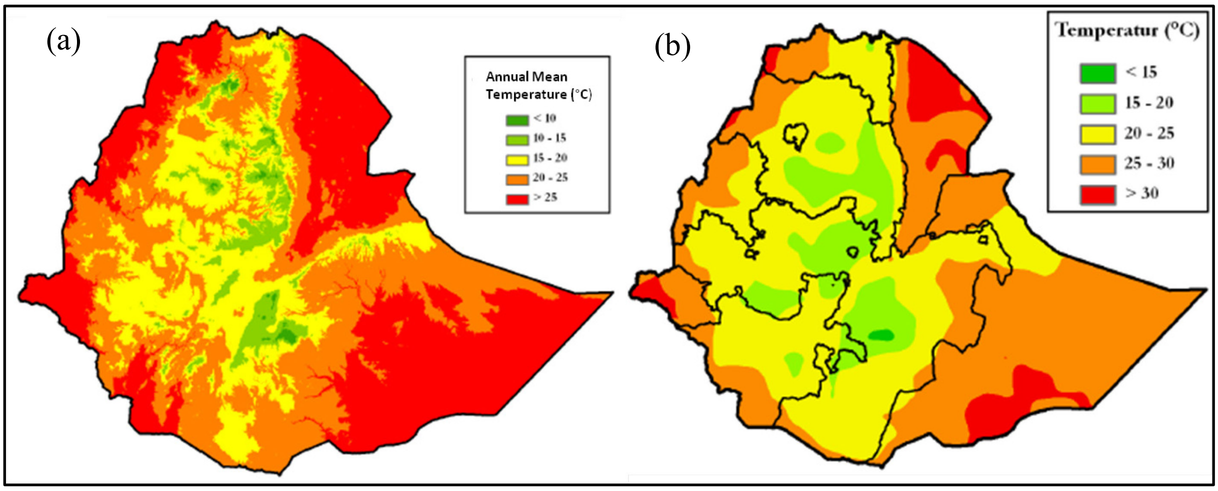
The Tyndall center for climate research provides several data sets of climate surfaces. The methodology for the creation of the climate change datasets is presented in (Mitchell et al.

2004). While originally developed at the Tyndall center the data is now provided and updated through the Climatic Research Unit (CRU) at the University of East Anglia, UK.

The entire dataset of the Tyndall center comprises both observed climate data and future projections at a resolution of 10 arc minutes for Europe and 0.5 arc degrees (30 arc min: about 50km depending on latitude) for the rest of the world. The generation of grid data is more complex than the rather straight-forward interpolation of data by WorldClim. A full description is documented in (Mitchell et al. 2004). Complexity is added by the inclusion of sunshine and cloud cover as variables while WorldClim uses only latitude, longitude and elevation as independents. The Tyndall TYN SC 2.0 data set that can be used for climate change projections includes monthly steps for precipitation, daily mean temperature, diurnal temperature variation, vapor pressure and cloud cover.

A comparison of the resulting temperature maps of the WorldClim data and the CRU data illustrates the differences in datasets for Ethiopia (Figure 6). The WorldClim dataset (Figure 6a) is much more detailed and gives a better representation of steep variation in mountainous areas than the CRU data (Figure 6b).

Figure 6. Comparison of WorldClim and CRU climate data



(a) WorldClim, (b) CRU interpolated; both maps show annual mean temperature in Ethiopia; most notably, a mountain range in the North appears to be misrepresented in the CRU set; maps adapted from Rüegsegger (2008).

2.1.1.4 Conclusion

Surfaces of historic climate data are an important step in climate change impact modeling. They provide the reference data for comparison of future states with current status. The quality of the data in the reference data is crucial for the quality of the resulting model.

Two datasets are most widely used; both have advantages and disadvantages for climate change impact modeling. The WorldClim data provides a basic set of 19 bioclimatic variables and monthly precipitation and temperature data that is based on an interpolation approach. Uncertainties are rather high in coffee regions and the available variables are limited. The CRU data offers a larger number of variables, especially also agronomically relevant data like radiation. However, uncertainties are not clearly documented, the resolution is coarser and the larger number of variables often limits climate impact studies to a reduced number of GCM outputs. Therefore, in following chapters the WorldClim data will be used as current climate data.

2.1.2 Generating future climate surfaces

Global circulation models are used to project future climatic developments. They are mathematical representations of the circulation of the earth's atmosphere. Radiative forcing models link circulation models with emission scenarios so that a projection of future climatic developments is computed. This data needs to be further processed to be useful for high resolution impact analysis. This climate modeling process is presented here, starting with the emission scenarios that drive the GCMs.

The IPCC report on the climate change that was most recent at the start of this thesis work was the 4th Assessment Report (AR4). In an attempt to make models more consistent the IPCC initiated the development of a set of emission scenarios based on demographics. The result was the Special Report on Emission Scenarios (SRES, (Nakicenovic and Swart 2000). The report groups possible future demographic developments and makes assumptions about the emission of these developments. These scenarios are commonly referred to as SRES.

The latest IPCC report, the 5th Assessment Report (AR5) updated the emission scenarios and now uses representative concentration pathways (RCPs). The concept is similar to the SRES but employs more updated data on the underlying future developments (Moss et al. 2010).

2.1.2.1 AR4 emission scenarios

The SRES scenarios are grouped according to storylines. The A1 scenario is further subdivided into 3 scenarios that are characterized by different energy sources. Key differences between the four storylines are assumptions about population development and economic development (Nakicenovic and Swart 2000):

- The A1 scenarios assume a rapid economic growth and a population that peaks in mid-century before it declines. Regions converge quickly, cultural and social interactions increase so that regional differences in income decline. The three sub-scenarios take different energy sources into account. A1FI denotes the intensive use of fossil sources, A1T non fossil sources and A1B a balanced set.
- A2 is characterized by a fragmentation of the world. Economic development does not converge. The population grows constantly until the end of the century.
- B1 like A1 is based on a converging world with a population that increases only until the mid of the century. The difference is the assumption of a change towards a service and information oriented economy with a reduced resource use.
- B2 is a regional world scenario like A2. Population grows constantly but slower than in A2; the technological change is slower than in B1 and more diverse than in A1.

For climate projections the scenarios A1B, A2 and B1 are most commonly used; B1 representing an optimistic case, A1B a moderate one and A2 as a scenario with high and increasing emissions.

The underlying assumptions of the scenarios about population and economic development are based on models themselves. The economic scenarios are based on a literature review. The authors of the report stress the difficulty to reliably assess future economic development, claiming that major determinants of long term economic growth, namely technological and institutional change, are exogenous to modeling. The estimates for future gross domestic product (GDP) growth within the SRES reflect the range of views within the literature (Nakicenovic and Swart 2000), Section 3.3.4). The projections about population development are based on a report by the United Nations. Assumptions about fertility rates, aging and urbanization yield estimates of population growth rates and associated emissions. Notably, the report rejects a strong causal association between GDP and population growth, seeing the two developments as largely independent of each other (Nakicenovic and Swart 2000), Section 3.2.5). The scenarios project population and GDP for the year 2050 as in Table 2.

Table 2. SRES projection economic and population scenarios

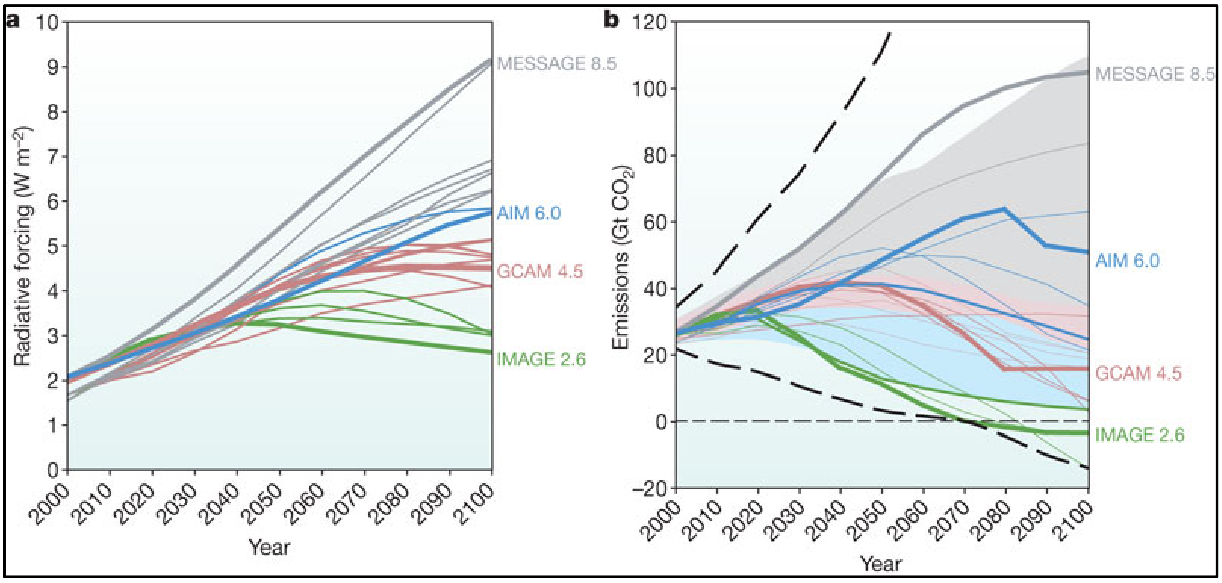
Scenario	A2	A1B	B2
Population in Billion (2050)	11.3	8.7	9.3
GDP (Trillion 1990 US$) (2050)	82	181	110

(Nakicenovic and Swart 2000)

2.1.2.2 AR5 emission scenarios

Global climate models from the 5th assessment report are driven by RCPs that describe radiative forcing scenarios in W/m² caused by greenhouse gas emissions (Figure 7) (Stocker et al. 2013). The RCPs were developed by collaboration of researchers of several fields. While harmonizing underlying data assumptions about population development, GDP development and carbon intensities differ (van Vuuren et al. 2011). The RCP 2.6 emission pathway assumes very low greenhouse gas emissions that peak before mid-century and then decline towards the end of the century (van Vuuren et al. 2007). In the RCP 4.5 scenario emissions are reduced and radiative forcing stabilizes until 2100 (Clarke et al. 2007). The RCP 6.0 is an intermediate scenario in which radiative forcing continues to increase until the end of the century (Fujino et al. 2006). RCP 8.5 is characterized by increasing GHG emissions over time resulting in high GHG concentrations (Riahi, Grübler and Nakicenovic 2007).

Figure 7. RCP emission pathways and radiative forcing

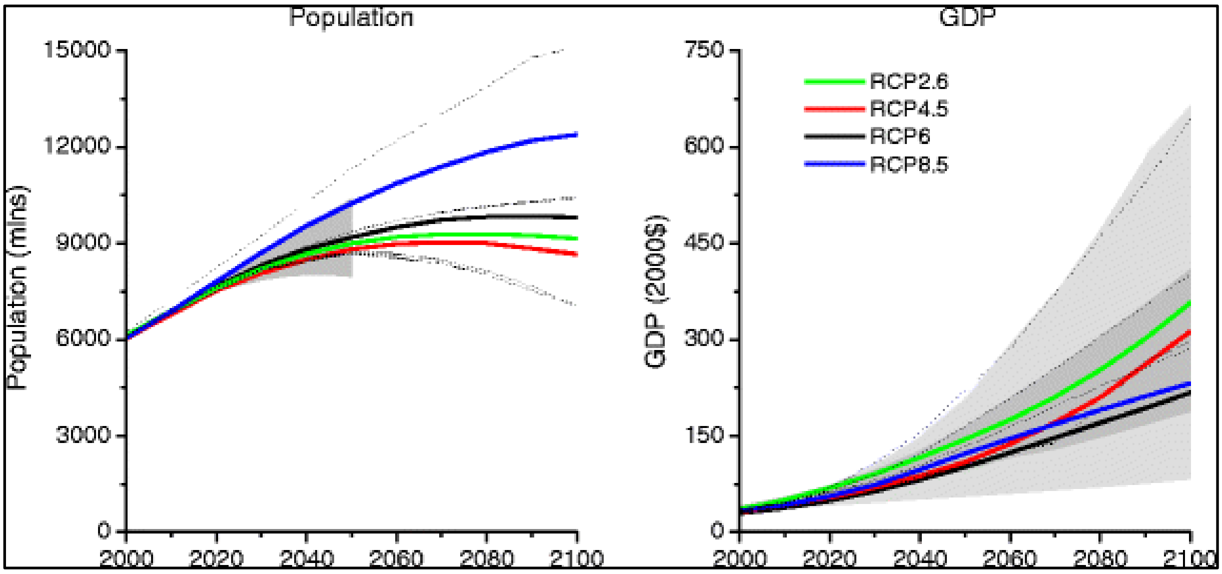


A) Changes in radiative forcing relative to pre-industrial conditions. Lines show individual scenarios from candidate RCP. B) CO2 emissions for the RCP candidates. Blue shaded area corresponds to mitigation scenarios; grey shaded area corresponds to reference scenarios; pink area represents the overlap between reference and mitigation scenarios (taken from Moss et al. 2010).

Of the four RCPs two have been criticized as unrealistic. RCP 8.5 is said to either assume too high oil consumption, or would result in too dramatic impacts (Inman 2011). RCP 2.6 was designed to demonstrate a feasible pathway towards a maximum of 2°C global warming. Such a target has recently been dismissed as unachievable (Frölicher, Winton and Sarmiento 2013).

The population and GDP pathways underlying the four RCPs are shown in Figure 8. Most scenarios assume a population by 2050 that is lower than in the SRES, but a higher GDP.

Figure 8. Population and GDP projections in the four RCP scenarios



Grey area for population indicates the range of the UN scenarios (low and high) of population development. Grey area for income indicates the 98th and 90th percentiles (light/dark grey) of the IPCC AR4 database. The dotted lines indicate four of the SRES marker scenarios (image taken from van Vuuren et al. 2011, p.17).

2.1.2.3 Global climate models

Climate models are constructed such that they incorporate fundamental physical laws, e.g. Newton's law of motion. Subsequently they are subjected to physical approximations of the climate system. Discretization of these approximations is necessary to limit computing time. This constrains the resolution of the models to grid cells of about 200km for AR4 models (Solomon et al. 2007), 8.1.3) and max. 100km for AR5 GCMs (Stocker et al. 2013). The models parameterize atmospheric processes, ocean processes, terrestrial processes, cryospheric processes, allow for atmospheric aerosol dynamics and couple the partial processes (Randall et al. 2007), 8.2).

The final models are very complex and consist of many different parts. The validity of a model and its components must be tested, i.e. the model must be compared with observations.

One way to do this is model intercomparison or ensembles. Intercomparison is conducted by running individual models several times to analyze internal variability. Furthermore the various models are compared by running unforced control simulations, simulations attempting to reproduce observed climate change over the instrumental period and simulations of future climate change. An ensemble may consist of individual models that are run to produce different versions by varying model parameters within plausible ranges. Another approach is to incorporate model results of different modeling centers.

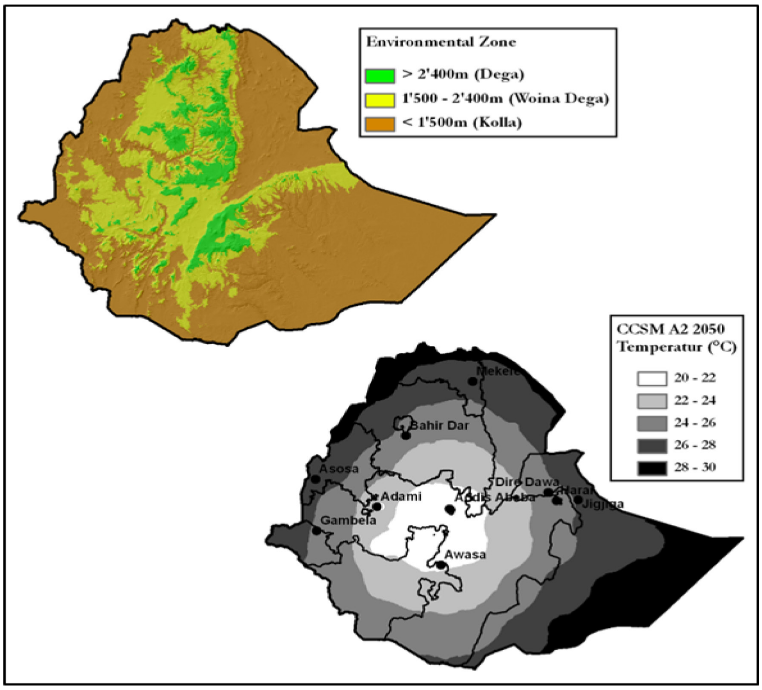
Model evaluation can also be done by analysis of metrics of model reliability. Metrics are chosen and correlated with model performance. That way, individual model projections are identified that are potentially more accurate. An important metric is climate sensitivity which means the change in global mean surface temperature following a doubling of atmospheric CO_2 concentration. Commonly model performance is tested by their ability to model past and present climate. The model projections are compared with observations to estimate the model accuracy (Randall et al. 2007), 8.1).

A critique to the conventional approaches to climate change impact assessment as discussed here has been published (Moss et al. 2010). The authors argue that until today research is sequential, i.e. the complex issue of climate change is disaggregated into a chain of causes and effect. Namely the chain starts at socio-economic scenarios that are the basis for emission scenarios. Radiative forcing scenarios then model atmospheric GHG cycles as a basis for climate model scenarios. They call for an integration of models to allow for feedback (e.g. between climate change impacts and socio-economic models) and better consistency. However, until better models exist impact assessment has to rely on the existing data.

2.1.2.4 Regional climate data

An example from (Rüegsegger 2008) exemplifies the need to provide more detailed climate data than provided by global models to model the impacts of climate change on agriculture. In this attempt to assess the climate change impact on coffee in Ethiopia the study first demonstrates the direct use of GCM data. Figure 9 shows the spatial distribution of environmental zones in Ethiopia in contrast with a GCM output for annual mean temperature (GCM: CCSM 3.0 NCAR, for 2050ies in the A2 scenario) for this region. The unprocessed projections would imply drastic changes of local climate conditions that appear to be unfeasible. For example local factors such as altitude are clearly not reflected. In conclusion, raw GCM outputs can be regarded as too inaccurate to reliably model impacts on local scale.

Figure 9. Environmental zones in Ethiopia and GCM temperature predictions



Environmental zones in Ethiopia by elevation gradients and CCSM 3.0 NCAR model predictions (A2/2050) for annual mean temperature for Ethiopia (taken from Rüegsegger 2008).

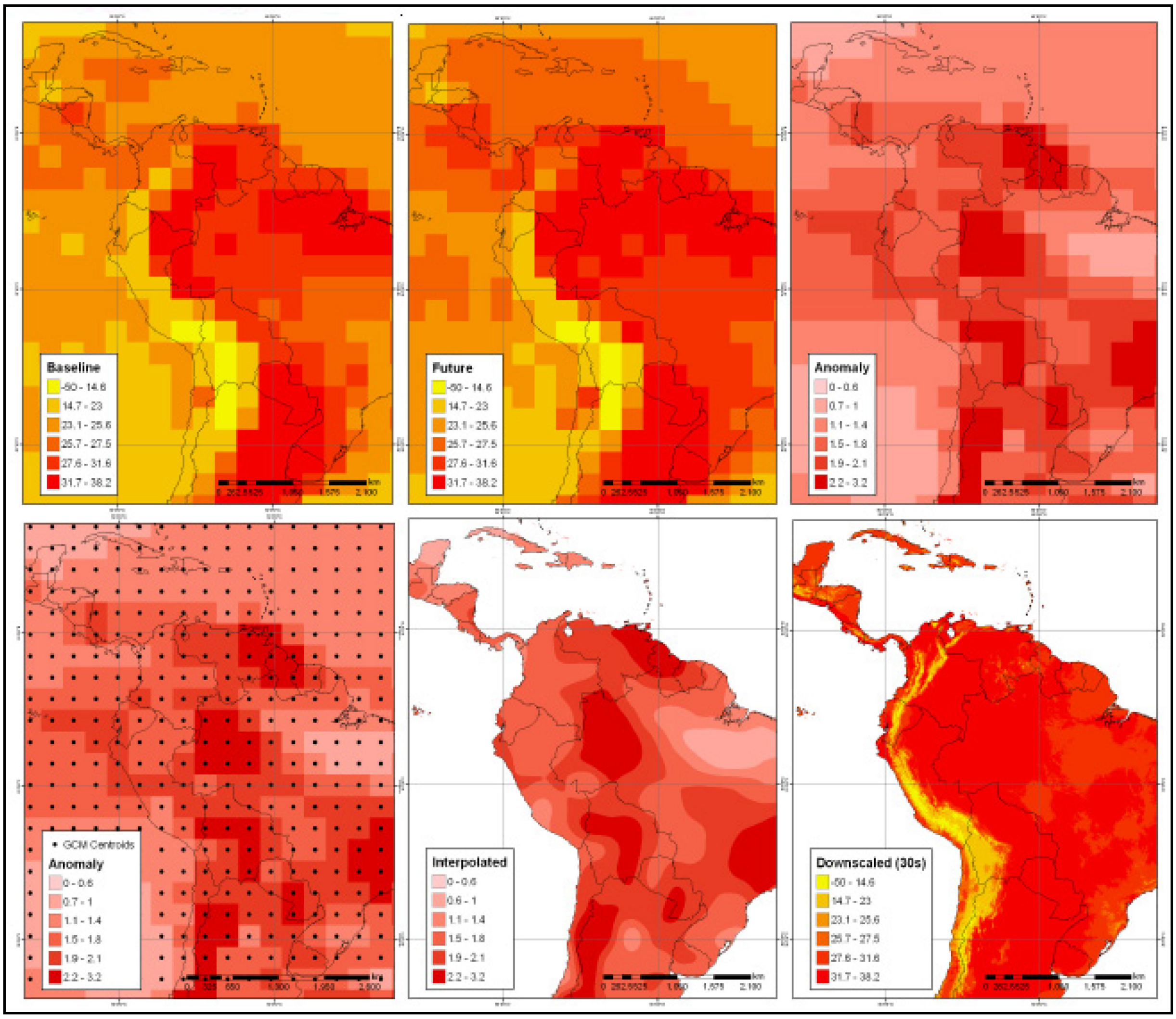
Different methods to provide spatially more detailed climate data exist. The development of regional climate data can alter results and is therefore an important step during the modeling of climate change impacts on crops. Basically, it is differentiated between dynamic and statistical downscaling. The former uses a nested regional climate model that uses GCM outputs as corner conditions for changes. The latter simply applies interpolated changes to current climate data. The advantage of this approach is the reduced computing time as compared to regional climate models like Precis. This makes it easier to provide climate data at high resolution for all GCMs and emission pathways of interest. Studies that use nested dynamic downscaling are often limited to single GCMs and do not account for disagreement between GCMs. Therefore statistical downscaling retains advantages.

A key underlying assumption of statistical downscaling is that climatic changes are relevant only at larger scales but that at regional scales the relationships between variables remain constant under future conditions. This assumption is taken to hold true for homogenous landscapes and considered to be problematic in more heterogeneous regions such as the Andes (Jarvis and Ramirez 2010). This way regional data sets for future scenarios are available for most IPCC GCMs and their data for the SRES and RCPs at several different 30-year running mean time slices (2020s to 2080s). The data is readily downloadable from (http://www.ccafs-climate.org/data/). At different spatial resolutions the same bioclimatic data as for the

WorldClim database is available: Monthly precipitation and mean, maximum and minimum temperature.

Statistical downscaling of GCM outputs as proposed by (Jarvis and Ramirez 2010) uses the WorldClim database as a baseline current climate reference. This surface is adjusted with changes in climate as projected by the GCMs. The method is named the delta method as it simply creates a smoothed surface of changes ("deltas" of climates and applies this to the WorldClim surface (Figure 10).

Figure 10. Illustration of the downscaling process of regional climate data



January maximum temperature in South America: (a) current climate data, (b) future climate data from GCM, (c) calculated climatic anomaly, (d) overlay with GCM centroids, (e) Smoothed interpolated anomaly surface, (f) combination of WorldClim current climate surface and downscaled future anomaly surface (taken from Jarvis and Ramirez 2010).

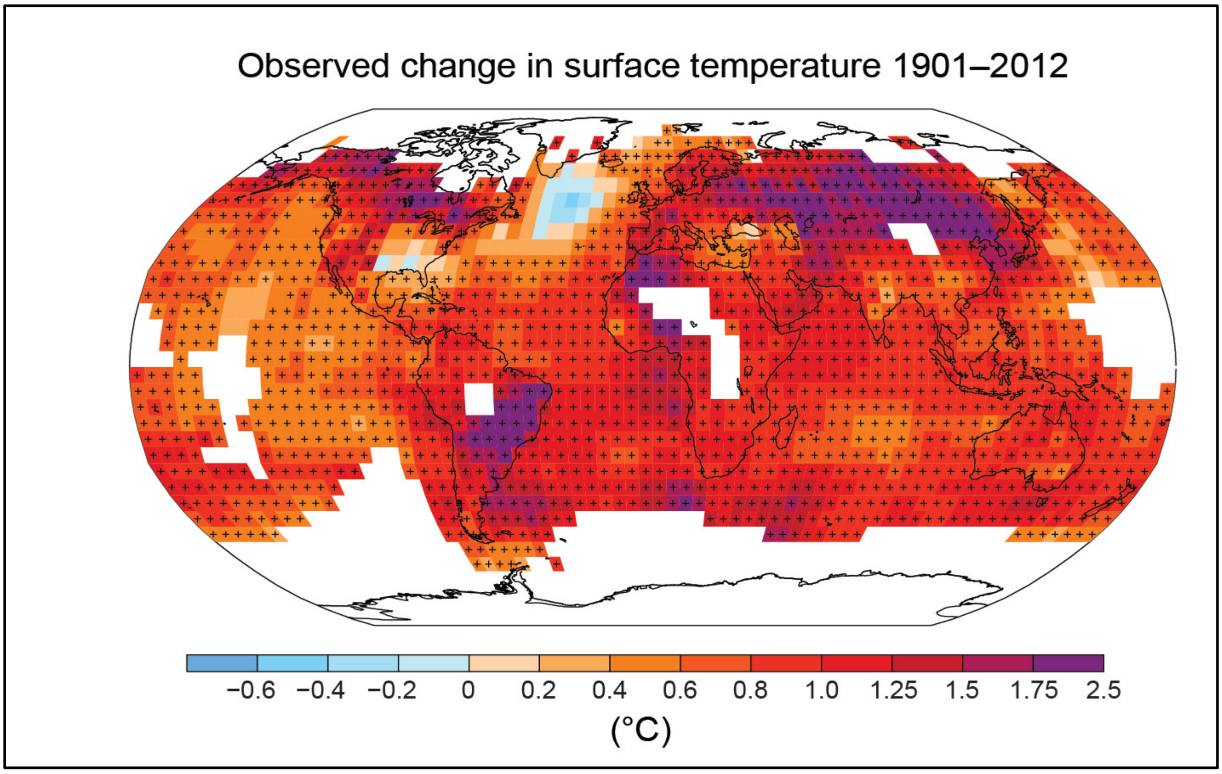
The procedure to produce smoothed climate surfaces of climatic changes resembles the one as described for WorldClim. Centroids of GCM cells are treated as reference points for interpolation of data. Average values of climatic changes are calculated from the GCM outputs for each reference point and time slice, and the thin plate spline interpolation algorithm is applied to yield a smoothed 30arc seconds surface. This surface is added to the WorldClim surface to yield a bias corrected surface for future climatic conditions (Figure 10).

2.1.3 Observed and projected climatic change

Since the last report of the Intergovernmental Panel on Climate Change (IPCC), the 5th Assessment Report (AR5) (Pachauri et al. 2014) climate change is no longer an academic exercise but has to be differentiated in to two aspects: observed changes to the climate, and future changes as projected by GCMs.

The observed changes can be attributed to human activity. The warming that is now unequivocal is often unprecedented in human history. Global average surface temperatures have increased by approximately 0.85°C over the past 120 years. These observed changes have not been distributed equally over the globe (Figure 11). Land masses have usually warmed more than oceans. From a coffee perspective the most notable change is the substantial temperature increase in Southern Brazil where today a third of all coffee worldwide is produced.

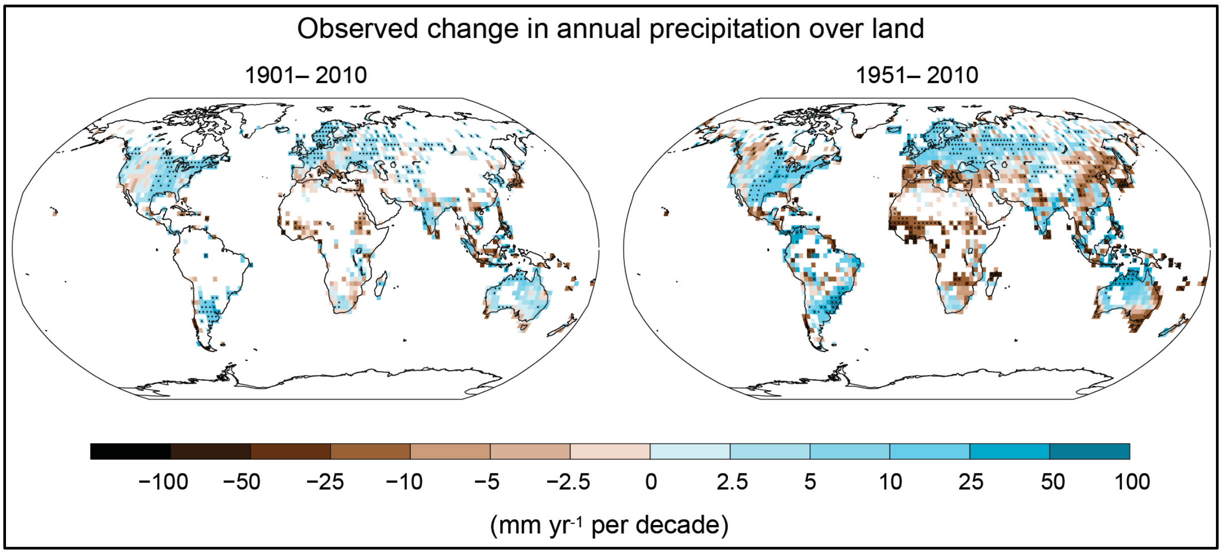
Figure 11. Observed change in surface temperature 1901-2012



Colored grid cells indicate trends where data availability permits a robust estimate; other areas are white (taken from Stocker et al. 2013).

Data for observed precipitation changes is less complete than the temperature data, especially when regarding accounts before 1951. No uniform trend could be observed. Some coffee regions have become drier, others wetter. The most obvious trend is that African Robusta and Arabica locations have become drier, especially West Africa and Ethiopia. For coffee in Asia and America trends were mixed (Figure 12).

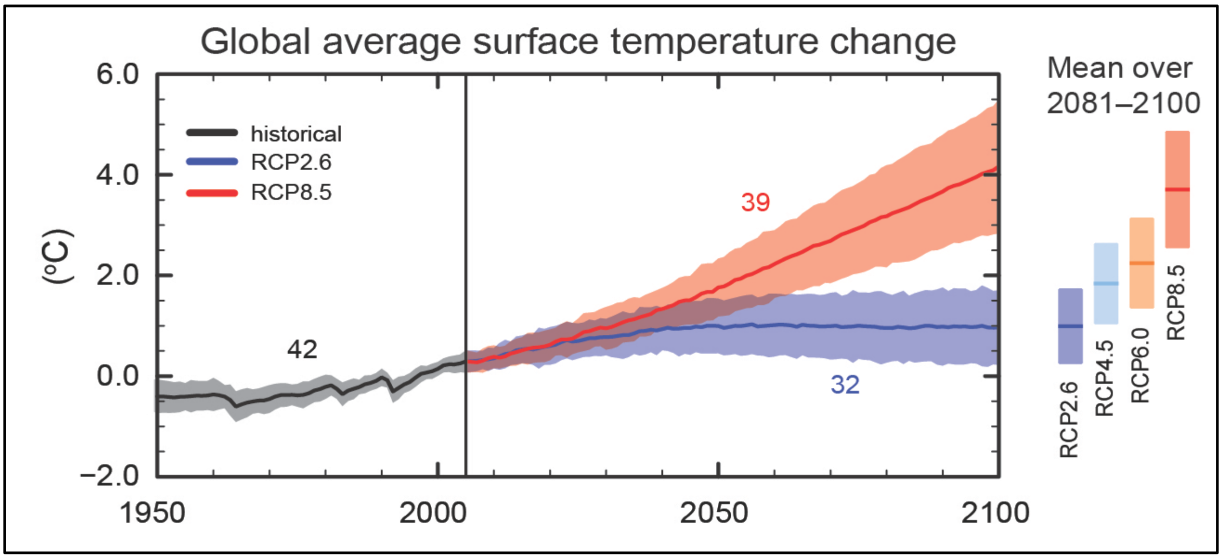
Figure 12. Observed precipitation change from 1901 to 2010 and from 1951 to 2010



Colored grid cells indicate trends where data availability permits a robust estimate; other areas are white (taken from Stocker et al. 2013).

Climatic changes in coming decades are projected by GCMs. Four different scenarios of anthropogenic GHG emissions (Representative Concentration Pathways - RCP) are used to project climatic changes. In the low emissions scenario RCP 2.6 global temperatures increase to a maximum of 2°C before stabilizing (Figure 13). With high emissions as in the RCP 8.5 scenario a continued temperature increase is projected of about 4°C until the end of the century.

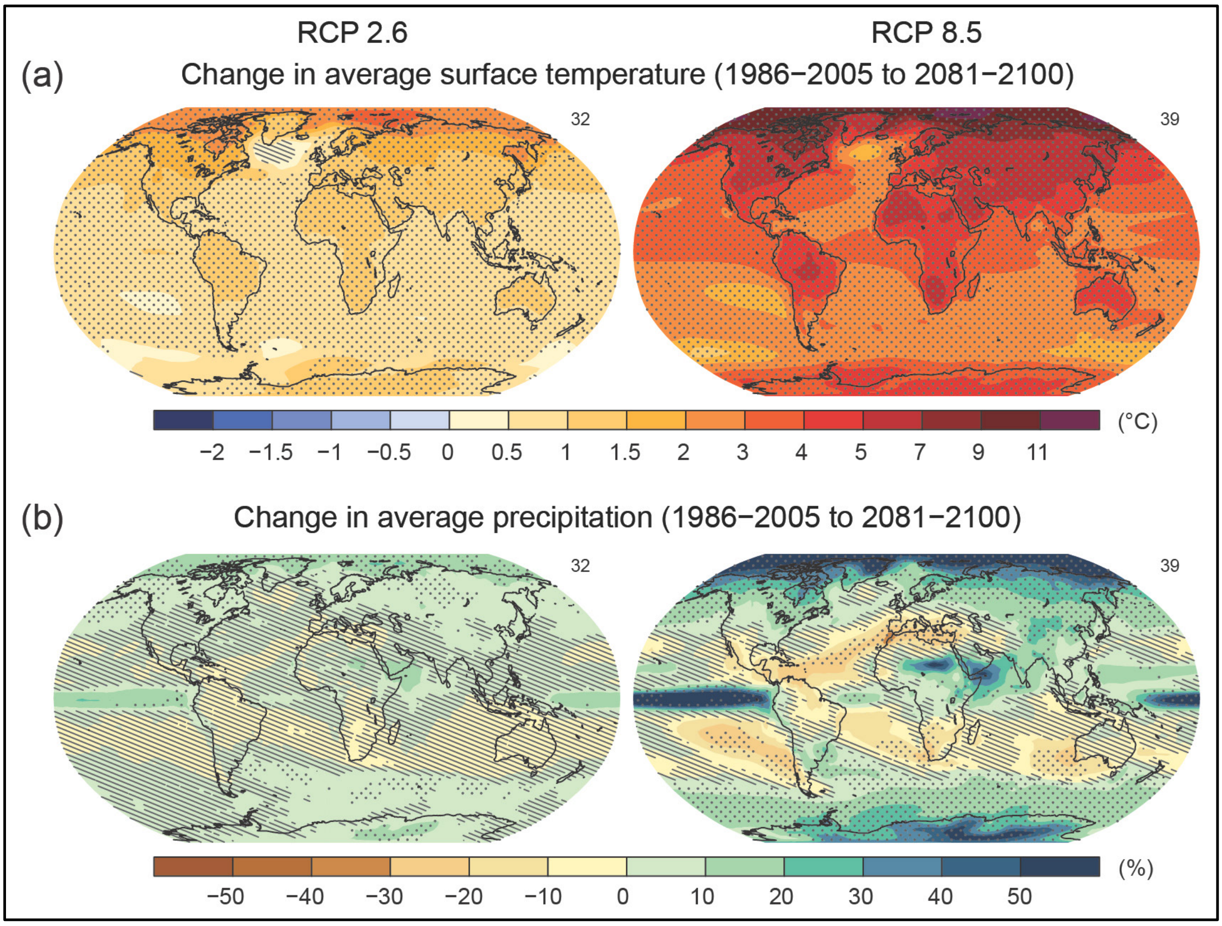
Figure 13. Projected changes in global average surface temperatures



Time series of projections and uncertainty (shading) for scenarios RCP2.6 (blue) and RCP8.5 (red). Black (grey shading) is the modelled historical evolution using historical reconstructed GHG levels; Mean and associated uncertainties averaged over 2081−2100 are given for all RCP scenarios as colored vertical bars; The number of models used to calculate the multi-model mean is indicated (taken from Stocker et al. 2013).

The impact of climate change on precipitation is not as unidirectional as the temperature increase. It is the precipitation gradient between dry and wet zones, and dry and wet season, that has been projected to increase (Stocker et al. 2013). Thus, dry zones become drier, wet zones wetter, dry seasons more pronounced with extreme precipitation events during the wet season. However, impacts will not be uniform across the globe. Like the observed early impacts (Figure 11) the coming impacts will also depend on the region. E.g. the arctic region will warm more rapidly than the rest of the globe. Precipitation changes in coffee regions will not be uniform. Some current growing regions will see precipitation increases; other decreases (Figure 14).

Figure 14. Projected climatic changes until 2100



(a) Average surface temperature and (b) average percent precipitation change in the RCP scenarios RCP 2.6 and RCP 8.5. Hatching indicates regions where the multi-model mean is small compared to natural internal variability; stippling indicates regions where the multi-model mean is large compared to natural internal variability and where at least 90% of models agree on the sign of change (taken from Stocker et al. 2013).

For temperature changes the modeling uncertainty is smaller than for precipitation. GCMs not only agree on the sign of change but also project that changes will be significant compared to natural variability. GCM uncertainty for precipitation is higher with less radiative forcing (RCP 2.6). The projected changes in this scenario are often within two standard deviations of natural variability. This is different with higher impacts in the RCP 8.5 scenario when projected changes will be beyond this threshold. However, there is little model agreement on the sign of change. Notably, in tropical regions, or the "coffee belt" between 30°N and 30°S model agreement is lower than in higher latitudes (Figure 14).

2.1.4 Conclusion

In the remainder of this thesis climate change scenarios form both the 4th and the 5th Assessment reports of the IPCC will be used. This was conditioned by the circumstance that the 5th report was published in 2013. Data from this report was made available for use between 2013 and 2014, so that some of this work had to be based on the earlier AR4 data. However, the improvement of GCM projections of the 4th Assessment report to GCM projections of the 5th Assessment Report is small for agriculturally relevant variables (Ramirez-Villegas, Challinor, et al. 2013) so that both datasets can be considered equally applicable. The data used here was downscaled using statistical methods. Dynamically downscaled climate data was not available at the time of writing.

Translating the change projections of climate science into scenarios of global change is the subject of widespread research itself. As agriculture is directly dependent on climatic conditions climate impact research is of broad interest in this sector. The first step in impact assessments in agriculture is the establishment of some kind of crop-climate relationship. Therefore the following sections will discuss this crucial input data.

2.2 Impact models for coffee

This section will elaborate the assumption that a changing climate may fundamentally change the coffee sector. First the climate dependency of coffee production is briefly discussed. Then, previous studies on the impacts of climate change on coffee production are reviewed. Emphasis in this review is put on methodological differences and results. A first rough distinction is made whether temporal variation is used to estimate yields or spatial variation to estimate spatial climate effects. The former models extrapolate historic climate variation on future scenarios. The latter models use differences in spatial variation to derive a function that describes the spatial climate dependence of coffee production. Most literature used the latter approach. These models will be reviewed in order of complexity.

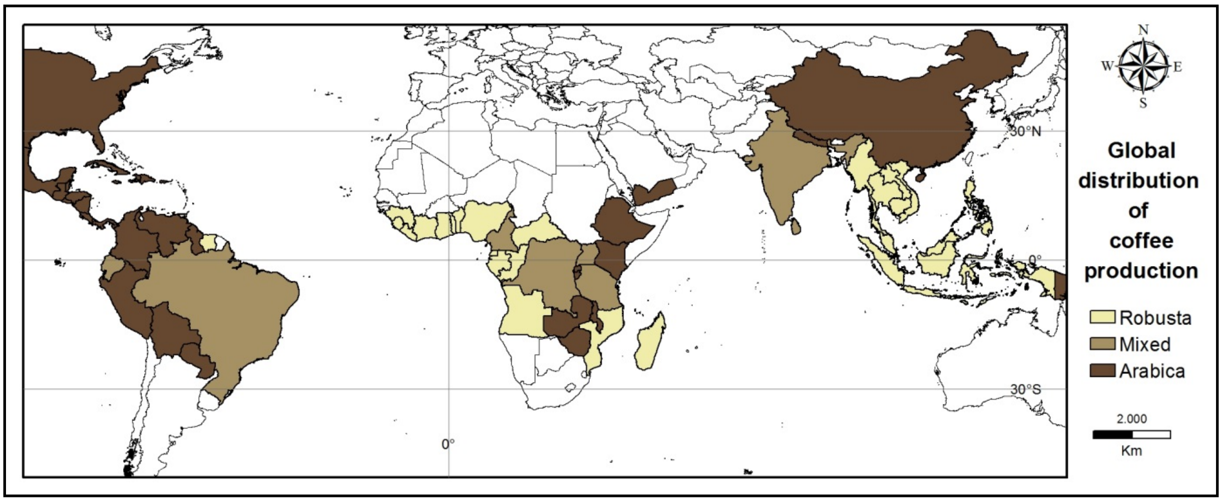
Usually research relied on a combination of temperature and precipitation variables. Although soil attributes, aspect, and local microclimate determine crop performance at local scales, they are unimportant in defining the global distribution. In some approaches annual mean climate values were used. Regional approaches often accounted for influential seasonal patterns in precipitation and temperature, e.g. dry season maximum temperature etc. In approaches that rely on methods developed for spatial ecology such concepts were generalized in the form of bioclimatic variables as described above (chapter 2.1.1.2). In a majority of studies changes in temperatures were found to have the largest impact.

Impact scenarios projected a migration of coffee production in altitude, or in latitude. The magnitude of impacts was often seen to be severe. Especially the studies that took into account long time horizons, e.g. 2080 or 2100, project a near complete loss of coffee production. However, with one exception all studies were regionally confined and no globally consistent modeling framework was applied to compare relative impacts across species and regions.

2.2.1 Climate dependency of coffee production

Botanically coffee belongs to the genus *Coffea* and comprises about 70 species. The main species that are used agronomically are *Coffea arabica* and *Coffea canephora* var. *Robusta*; minor ones are *Coffea liberica* and *Coffea excelsa* which account for only 1-2% of global production. All species have their origins on the African continent. *C. arabica* comes from Ethiopian high plateau areas and is naturally found between 1300 and 2000m. *C. canephora*'s natural habitats are found below 1000m in tropical Africa. *C. liberica* and *C. excelsa* originate from lowlands in Western and Central Africa. Coffee is grown in tropical countries along the equator between 22°N and 26°S (Figure 15). It takes approximately three years from germination to first fruit production. The shrub then remains productive for up to 80 years, though the economic lifespan is maximum 30 years (Wintgens 2009), p.4).

Figure 15. Global distribution of coffee production



Average area shares for each coffee species '98-'02; (own representation of data compiled from FAO 2012; USDA 2012; ICO 2013 as described in chapter 4).

The key environmental factors that influence coffee productivity are temperature, water availability, sunshine intensity, wind, type of soil and topography of land. The optimal mean temperature for *C. Arabica* is considered to be 18°C during the night and 22°C during day time. Extremes should not be lower than 15°C during night and not exceed 25-30°C at daytime. Reduced photosynthesis at temperatures above 25°C and a loss of flowers or fruit degeneration at temperatures above 30°C compromise productivity. Low temperatures favor diseases. Temperatures lower than minus 2°C for more than 6h are potentially lethal for the plant. *C. canephora* var. Robusta is generally more tolerant towards high temperatures but may die at 4-5°C.

Arabica requires about 1400 to 2000mm (min 800-1000) of annual rainfall, Robusta between 2000 and 2500mm (min 1200). Values lower than the minimum are potentially damaging for production. Excessive rainfalls are mostly a problem because of top soil erosion. A dry season of about 3 months is considered to promote productivity. Atmospheric humidity has an influence on transpiration and is therefore linked with necessary rainfalls. Ideal humidity is 60% for Arabica and 70% for Robusta (Descroix and Snoeck 2009), p.168). Recent publications demonstrate, however, that such general recommendations are overly generalized. A differentiation of coffee growing sites according to characteristics like slope, shade level, variety and others shows that management recommendations need to be site specific (Läderach, Oberthür, et al. 2011).

Wild coffee naturally grows in the lower levels of forests. Traditional coffee cultivation therefore uses shaded cultivation. Research has shown that extensive sunlight is counterproductive to photosynthetic efficiency. However, as only the upper leaves are fully exposed no-shade cultivation is possible. In such a system the plant requires additional nutrients and a more intensive management. This allows for a higher productive capacity. Without careful cultivation without shading the plants may exhaust rapidly and productivity drops. Extreme temperatures cause leaf burn or frost damage. Temperatures below 10°C damage chlorophyll levels and compromise photosynthesis. Shade trees reduce this risk. Other advantages of shade production are the reduction of erosion, additional organic matter in the soil, reduced weed growth, a curbing of the biennal bearing[1], improved bean size and aromatic quality (Wintgens 2009), p.20).

While light winds can be beneficial for plantations, strong winds may cause serious damages. Many important coffee growing areas are situated in regions that are prone to tornados or cyclones, e.g. Vietnam or the Caribbean. Constant winds like sea breezes may necessitate windbreaks.

[1] Biennal bearing describes the phenomenon of significantly higher yields in alternating years

Top soil should be at least 2m deep because of the deep root system of coffee plants. A constant water supply needs to be ensured. The soil should therefore have good water retention especially in dryer areas. A high water table or flooding may easily kill plantations by asphyxiation. The topography of plantations is ideally flat or slightly rolling because deep soils are offered. Coffee can be grown on steep slopes but this requires conservative methods to prevent erosion.

2.2.2 Yield estimation models

At the time of writing three models that attempt to estimate yields based on historical time series had been published. The most widely cited one is a paper by (Gay Garcia et al. 2006) on Mexican coffee production that uses a multiple regression approach. A similar method has been employed for Tanzania (Craparo et al. 2015). The most complex approach is the Caf2007 virtual *C. arabica* process model by Van Oijen et al. (2010b).

2.2.2.1 Veracruz, Mexico

The publication by (Gay Garcia et al. 2006) attempts to assess the impacts of climate change on coffee production in Veracruz, Mexico, by employing a regression model. The model includes economic and climatic factors to gain information about possible changes. The available climate data was then analyzed for trends and extrapolated into the future to estimate climate change impacts. First a general regression model was defined before eliminating insignificant variables using a backward stepwise elimination approach. Variables were removed from the equation if they did not contribute positively to the adjusted R^2 value. The final model is estimated as follows (Eq. 1):

$$\begin{aligned} Pcoffee = & -35965262 + 2296270\,(Tsumm) - 46298.67\,(Tsumm)^2 \\ & + 658.01618\,(Pspr) + 813976.3\,(Twin) \\ & - 20318.27\,(Twin)^2 - 3549.71\,(MINWAGE) \end{aligned} \quad \text{(Eq. 1)}$$

where:

P_{coffee} is the production of coffee.
T_{summ} is the average temperature during June, July, August.
P_{spr} is the average precipitation during March, April, May.
T_{win} is the average temperature during December, January, February.
MINWAGE is real minimum wage.

The adjusted R^2 value of the model is 0,692. Based on this regression model implications for climate change scenarios are derived. Holding other variables constant, an increase of two standard deviations, or 1.5°C, of summer temperature would cause a 30% yield loss. The respective value for winter temperature is 1.6°C and would cause a 50% yield loss.

Based on these values economic implications were discussed. Assuming constant costs the authors argued that revenue from future yield would not be sufficient to generate profits, and that coffee production could be abandoned in this region.

2.2.2.2 Tanzania

(Craparo et al. 2015) used data from several coffee plantations for plot scale models, and official survey data to estimate climate effects on yield in the northern Tanzanian region. From local climate data indices of bioclimatic variables were constructed and tested for their influence on observed yields in the study region.

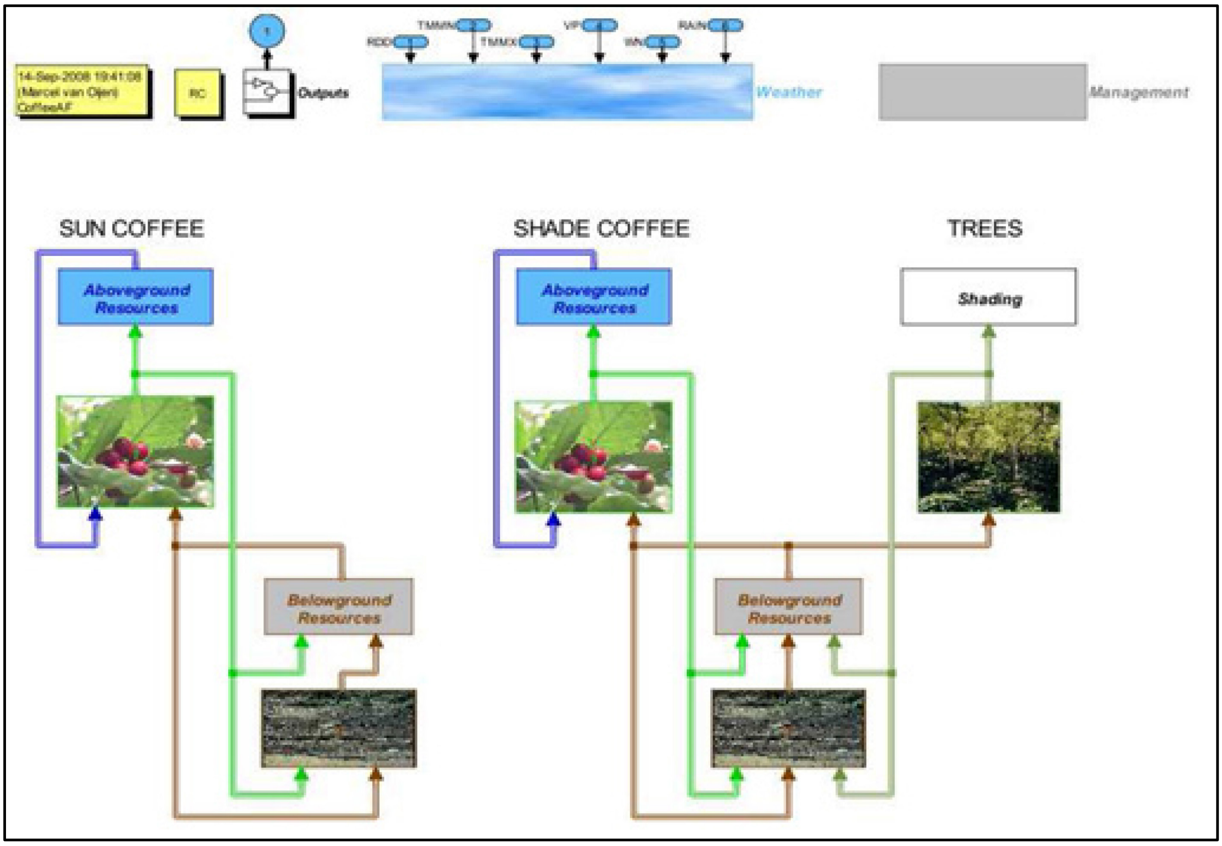
Increasing night time temperatures were found to have reduced yields over the past decades. The regression model showed that a 1°C increase of this temperature correlated with a 137 kg/ha loss of yield. Without adaptation their model projected yields of about 145 kg/ha by 2060. The study concludes that such changes have already reduced yields in the past and in the future might therefore have a substantially negative yield effect by causing sporadic and incomplete flowering periods.

2.2.2.3 Caf2007

A dynamic process based crop growth model of *C. arabica* has been developed at CATIE (Centro Agronómico Tropical de Investigación y Enseñanza) (Van Oijen et al. 2010b). Calibrated for the Turrialba region in Costa Rica it is able to reproduce historic yield developments despite an uncertain foundation of quantitative data.

The aim was to develop a model that takes into account the specific agronomic practices of coffee production. Namely, full plantation life cycles of 10-25 years and agroforestry shaded production can be simulated. Competition for light, water and nutrients was incorporated and for several management practices trade-offs between diverging goals can be evaluated. The conceptual structure is shown in Figure 16.

Figure 16. Conceptual representation of the CAF2007 dynamic process model for coffee agroforestry systems



Provided by O. Ovalle-Rivera, CATIE.

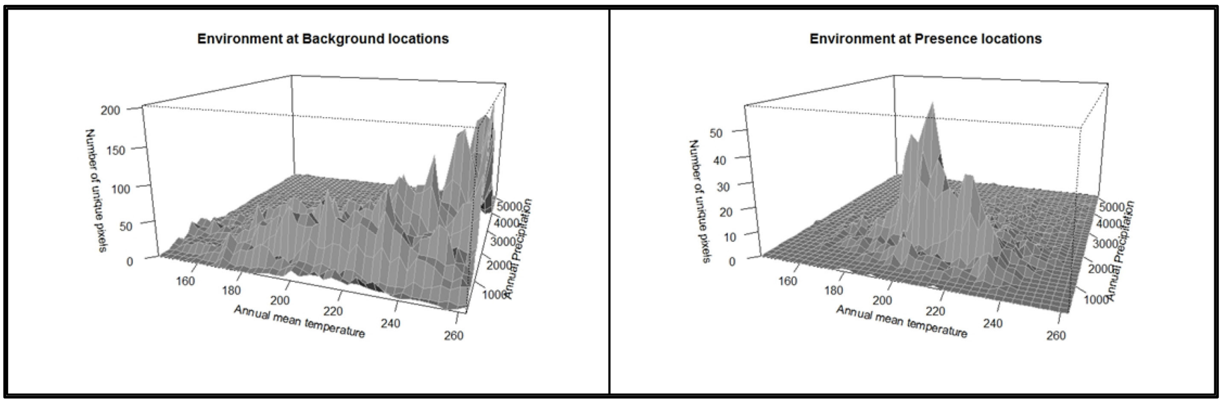
To demonstrate the applicability of the model the potential impacts of a 5°C warming in the Turrialba region were assessed (Van Oijen et al. 2010b). The authors concluded that increases in atmospheric CO_2 levels may increase the effectiveness of nitrogen fertilization. Nevertheless, the warming was projected to significantly decrease yields of coffee trees. In addition, warming could hinder the growth of shade trees for wood production (Van Oijen et al. 2010b).

Currently efforts are underway to improve model performance and calibration for other regions. Additionally, a parallelization of the model in order to model a region is planned. These two developments would make the model suitable for improved climate change impact assessments for coffee that explicitly include progressive adaptation.

2.2.3 Spatial distribution models

Rather than using time series to extrapolate past yield developments on future climate conditions locally, spatial approaches use indicators of mean climate conditions. Indices of climatic suitability are estimated based on some form of geographic occurrence data of coffee production and background data from the general environment (Figure 17). Some authors have relied on simple descriptive functions based on annual means; others fit more complex models using several independent variables using machine learning approaches. The origins of the approaches used for coffee come from both agriculture and also ecology. A key difference is that in agriculture environmental effects on the crop physiology are typically known, while in contrast they are unknown in ecology. As the quantitative knowledge on crop physiology is sparse for coffee (Van Oijen et al. 2010a) both approaches have been used for coffee.

Figure 17. Comparison of the environment at background and occurrence locations in spatial distribution models



Distribution of climate values (annual mean temperature in 0.1°Cand total precipitation in mm) at (A) background locations randomly sampled from coffee countries (B) *C. arabica* occurrence locations (own data and representation).

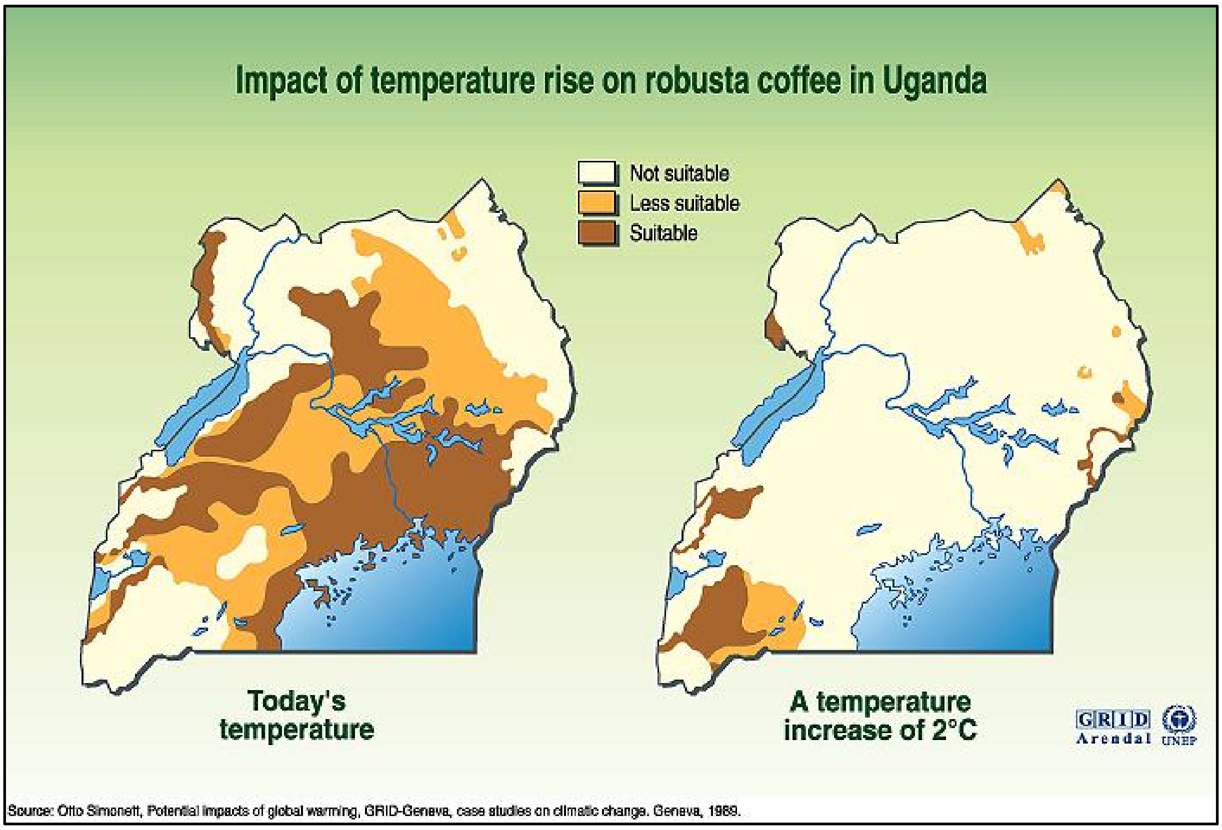
The best known approach from agricultural sciences is of the Agro Ecological Zoning (AEZ) type. Spatial modeling of this kind makes use of expert knowledge on the physiological limits of a crop to describe the suitability of a location for a crop. A limited number of variables is used and suitability functions are often easily comprehendible as they describe one-sided distributions. Spatial occurrence data is used for model assessment. The ecology derived approaches on the other hand often do not require previous knowledge on physiological limits. Instead of a large set of variables a probability function is fitted that separates sites of validated occurrences from sites of putative absence. The assumption is that a species can be found only at locations that show suitable climatic conditions.

The following is a review of literature sorted by research groups. The earliest research assessed the impact of climate change on Ugandan Robusta production only using annual mean temperature. A study on Ethiopia added annual mean precipitation. Several studies used a largely identical approach to assess the impacts of climate change on Brazilian coffee production but adding frost risk. The same AEZ type of modeling, but with much higher complexity is the global agro-ecological zoning (GAEZ) project. At the International Center for Tropical Agriculture (CIAT) three different attempts were made to improve impact models for tropical crops. The EcoCrop model resembles the GAEZ model. CaNaSTA and Maxent are spatial regression type approaches that have a background in spatial ecology.

2.2.3.1 Uganda

The earliest example of an impact study that assesses the effects of climate change on *C. canephora* production in Uganda is (Simonett 1988). This study is widely cited, but the original publication cannot be found. It can only be speculated about its methodology, inferring from a graphic (Figure 18).

Figure 18. The impact of rising mean temperatures on Robusta in Uganda



(Taken from Simonett 1988).

From shape and appearance of the map it can be speculated that it was based on interpolated temperature data for current conditions in combination with a rule that in Uganda Robusta varieties grow best at annual mean temperatures of 22°C to 26°C. The future scenario would then be derived by adding 2°C to the mean annual temperature. As Figure 18 shows area for Robusta cultivation would be drastically reduced and only remain feasible in mountainous areas.

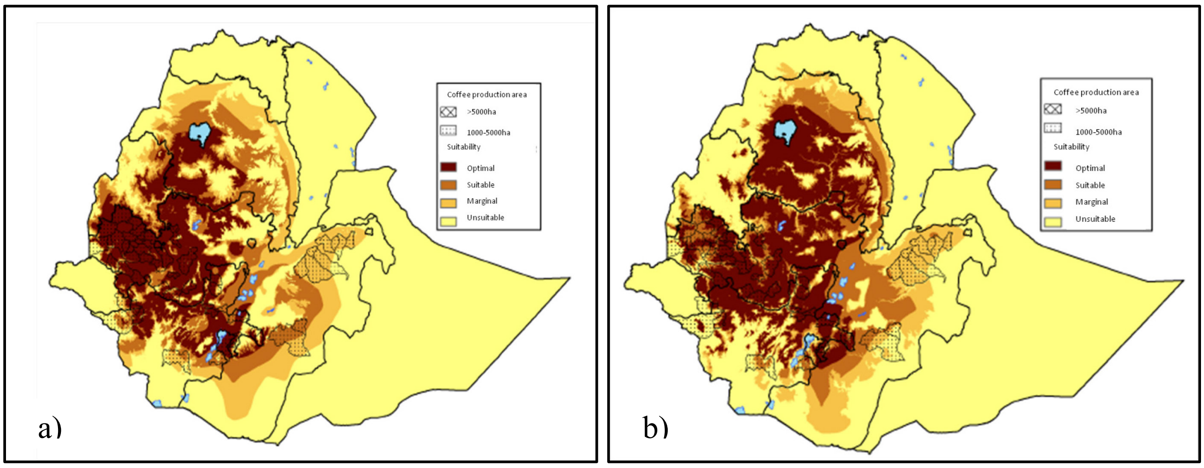
2.2.3.2 Ethiopia

A study by (Rüegsegger 2008) used distribution maps of known coffee growing areas in Ethiopia to calibrate a Fuzzy Logic based model. Fuzzy Logic is a method to convert binary problems into a non-monotonic ramp-like function. Similar to the Ugandan research approach values of annual mean temperatures were used, but also a precipitation variable was added.

In a repeated modeling exercise different thresholds and fuzzification parameters were tested about their power to correctly predict current growing areas. Initial values taken from the literature to define the lower threshold for annual mean temperature were between 15°C and 18°C, and the upper threshold to be between 21°C and 25°C. For precipitation only a lower threshold was defined to be between 1000 and 1300mm/year. After the modeling exercise these ranges were redefined as follows: Optimal temperatures between 17°C and 23°C , lower and upper threshold values 15°C and 25°C; Optimal precipitation 1200mm/year, and minimum precipitation 800mm/year. Sites with values that fulfill optimal conditions were labeled “optimal”, sites with values below minimum conditions “unsuitable”. Sites that have conditions within ranges are classified between “0” (unsuitable) and “1” (optimal).

Here two examples from the results sections are shown: A map of current suitability for coffee production with known areas of production and a map of projected future suitability in 2050 under the A2 scenario (Figure 19).

Figure 19. Ethiopia current and future suitability in a fuzzy logic envelope model



a) Suitability in current conditions; b) Suitability in the A2 scenario, 2050s. Black lines represent provinces, dotted areas represent known coffee production areas, dark colors indicate optimal conditions, light yellow unsuitable area (taken from Rüegsegger 2008)

The study concluded that 70% of optimal growing areas will lose suitability. Novel areas further north may replace lost acreage. The change was driven by rises in temperatures. Changes in precipitation were seen to have little influence on the result (Rüegsegger 2008).

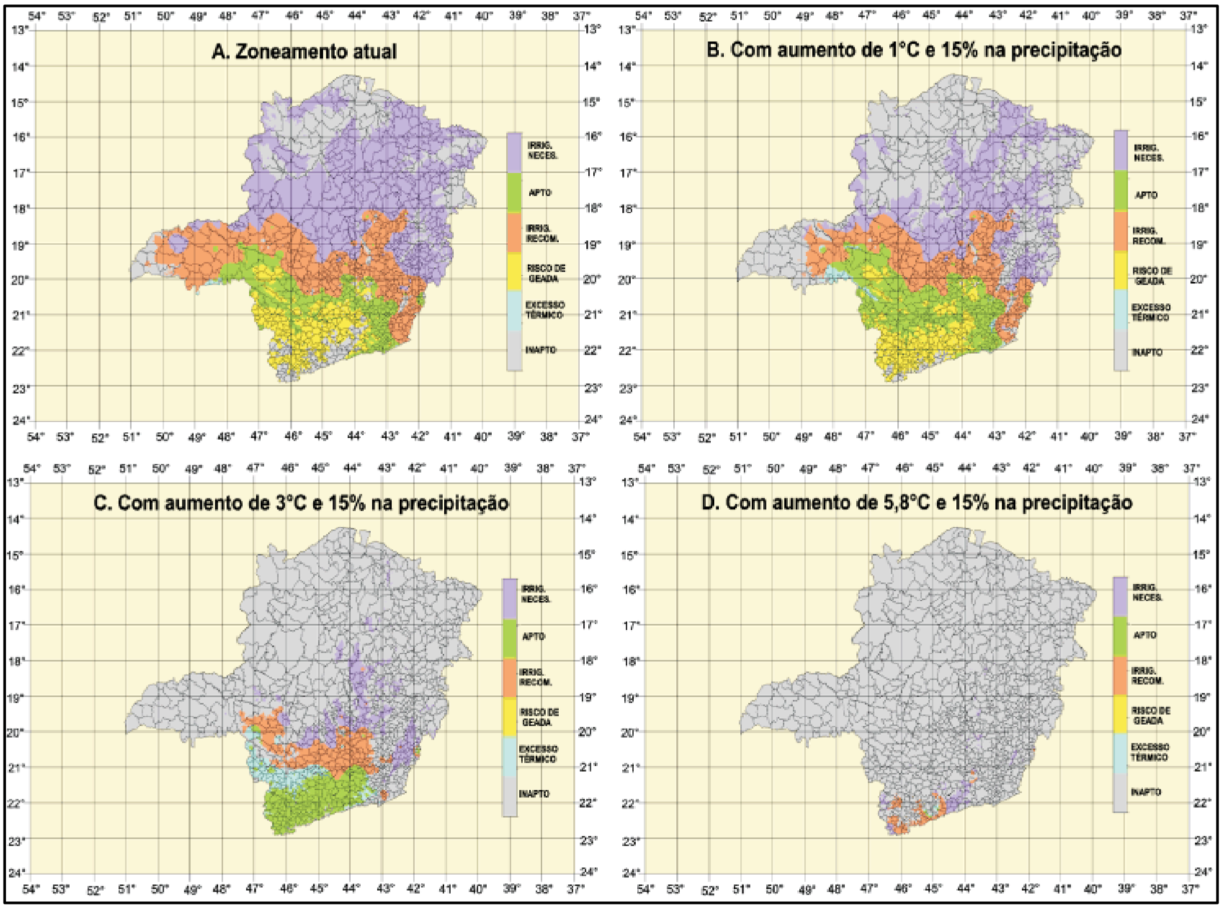
2.2.3.3 Brazil

The Brazilian research team employed an agricultural zoning approach to assess the impact of a changing climate on coffee production in Brazil. Based on an existing zoning climate parameters were altered to infer the impact of increased temperatures and precipitation as projected by the IPCC reports for Brazil. Initial publications used an approach that incorporated hypothetical increases in mean annual temperature by 1, 3, or 5.8°C (Assad et al. 2004; Zullo Jr, Pinto and Assad 2007; Zullo Jr et al. 2008). The latest publications involved the application of the Precis regional climate model (Assad and Pinto 2008; Zullo et al. 2011).

According to the authors the work goes back to a detailed agricultural zoning program initiated by the Brazilian Ministry of Agriculture following a series of severe harvest losses in the middle of the ‘90ies. The expectation had been to reduce losses by zoning according to precipitation and temperature needs of main cultivars. Thus, cultivar specific zoning approaches were used. For coffee the zoning was based on climate requirements that were proposed by Brazilian experts. The model assumes that coffee ideally grows in regions with (i) an annual water deficit of 0 to 100mm, (ii) average annual temperatures between 18°C and 22°C, and (iii) a frost risk of less than 25%. Areas with annual mean temperature between 22°C and 23°C and a water deficit up to 150mm were considered suboptimal. Frost risk was considered the strictest criterion (Zullo Jr et al. 2008).

The Brazilian researchers produced several publications concerning the impact of climate change on coffee production. Here, an example from (Assad et al. 2004) is shown (Figure 20). It shows suitability for coffee production in Brazil’s most important coffee state Minas Gerais under different future scenarios. For this publication different scenarios about future changes were developed from GCM data, e.g. an increase of 1°C in annual mean temperature.

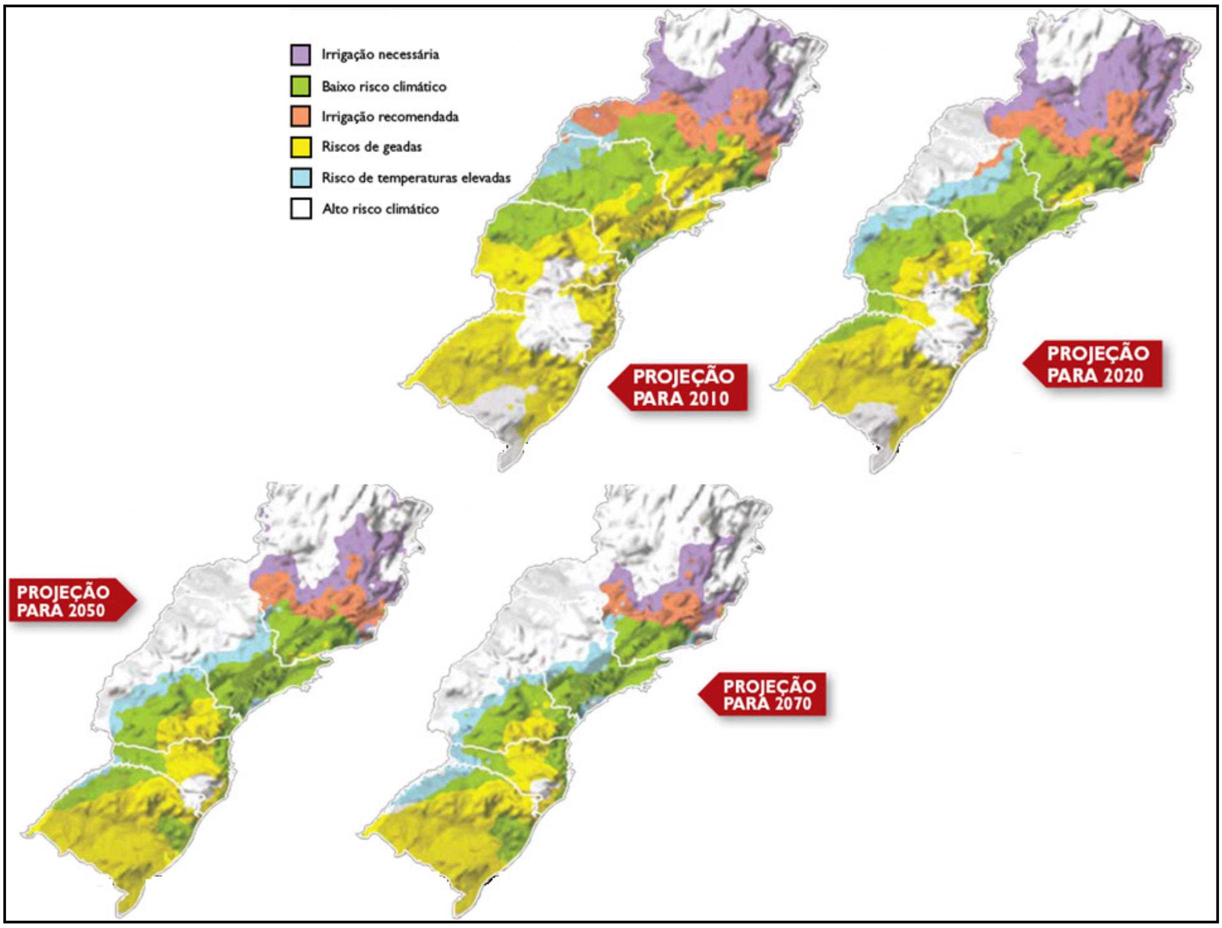
Figure 20. Climate change impact in Minas Gerais state (Brazil)



(A) Current suitability, (B) 1°C temperature increase and 15% more precipitation, (C) increases of 3°C temperature and 15% precipitation, (D) 15% precipitation increase and 5.8°C temperature increase. Zones shown in violet require irrigation, green have low climatic risk, for red zones irrigation is recommended, blue zones have a risk of high temperatures, yellow zones have frost risk; white zones are unsuitable (taken from Assad et al. 2004).

In later work the Precis regional climate change model was applied for the B2 and A2 scenarios of the HadCM3 GCM. This scenario projected a 2°C to 5.4°C temperature increase in the more severe scenario and a 1.4°C to 3.8°C temperature increase in the more optimistic scenario for Brazil by 2100. No results about changes in precipitation were reported in the publication (Assad and Pinto 2008). Maps for the distribution of risk zones in Brazil resulted (Figure 21).

Figure 21. Suitability for coffee production in important Brazilian coffee regions



Projections until 2070, A2 Scenario; zones shown in violet require irrigation, green have low climatic risk, for red zones irrigation is recommended, blue zones have a risk of high temperatures, yellow zones have frost risk; white zones are unsuitable (taken from Assad and Pinto 2008).

Relating these impacts to the number of municipalities that are suitable for Arabica production it is shown that in the Northern Minas Gerais region only a third will remain productive by 2070. In the South on the other hand several municipalities could enter production (Assad and Pinto 2008). Thus, this simple model projected a strong southward migration of coffee production in Brazil. Especially Northern areas with little precipitation were projected to be lost despite irrigation.

2.2.3.4 Global agro-ecological zoning

To evaluate the global potential for the production of several major crops based on climate and soil data is the objective of the global agro-ecological zoning project (GAEZ). Its latest version GAEZ v3.0 was developed by the International Institute of Applied Systems Analysis (IIASA) in Austria. It goes back to previous AEZ efforts by the Food and Agriculture Organization of the United Nations (FAO) as early as in the 70ies that were improved based on novel data and knowledge (Fischer et al. 2012).

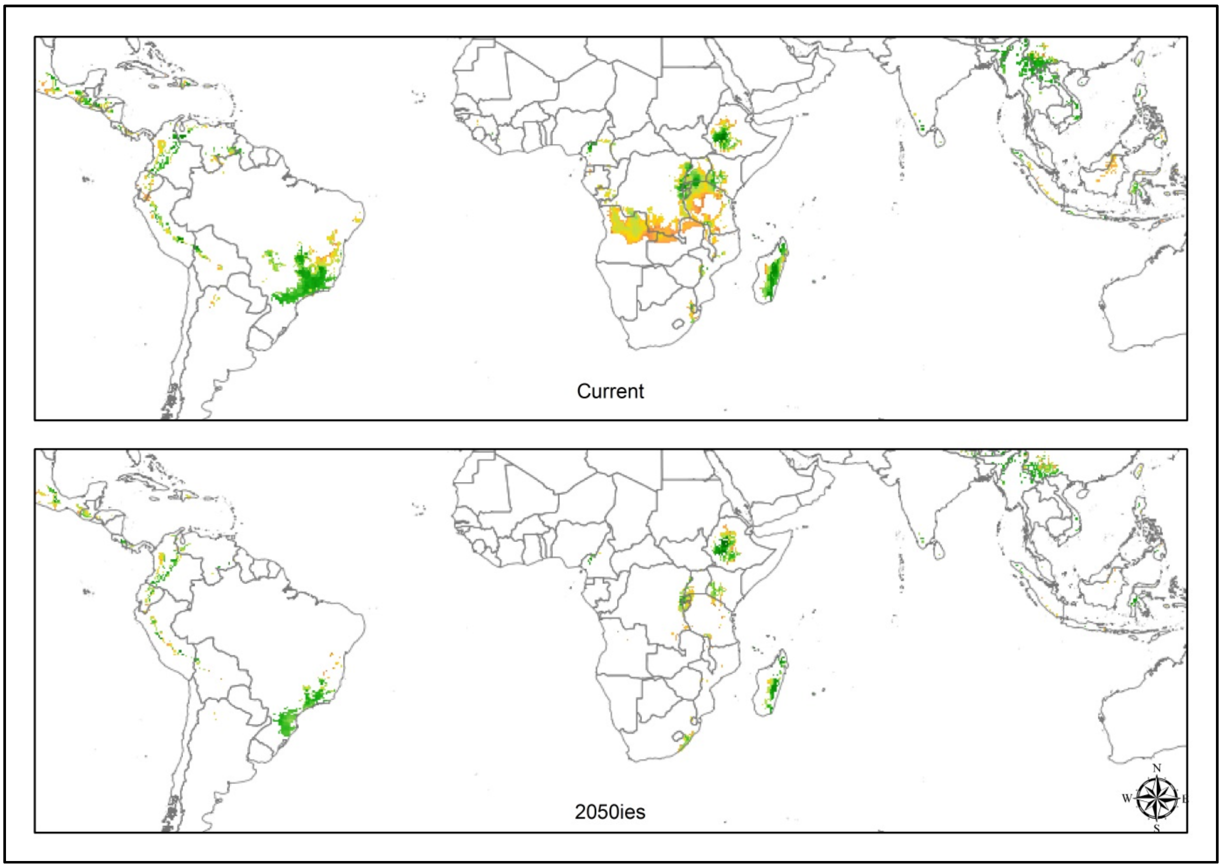
The approach represents a very elaborate agro-ecological zoning model. In an initial step monthly climate indicators that are meaningful for agriculture are compiled. The second step defines maximum attainable yields based on land use type specific crop calendars for each grid cell. These yield levels are corrected by applying climatic constraints in the third step, indicating agro-climatic attainable yields. In steps 4 and 5 this information is supplemented with soil and topographical constraints to generate global suitability maps. Subsequent modules make use of this information to estimate actual yield, yield gap, harvested area and production by grid cell. For four AR4 GCMs and three time steps the entire model is re-run to generate future impact scenarios (Fischer et al. 2012).

The estimation of agro-climatic suitability comprises two steps: First, the calculation of maximum attainable yield by land use type and next the correction with constraint factors. The latter step attempts to account for climate-related constraints such as pests and diseases. Thus, only the first of these two steps is directly related to the specific crop. For each crop for a number of variables functions define whether a grid cell is optimally suitable, sub-optimal or unsuitable under absolute constraints. Variables include temperature, precipitation, evapotranspiration and radiation variables. The process optimizes the growing period, estimates effects of water deficit, and CO_2 fertilization. Some functions are two sided constraints, others are one sided. Constraints based on water deficit are only applied to land use types that are rain-fed but not to irrigated land (Fischer et al. 2012).

It is difficult to evaluate the exact relationship for a single crop in GAEZ as factors, limits, and functions and their respective justification with other literature are spread out over several chapters, annexes and additional reports, or are unattainable. For example, maximum attainable yield levels for coffee appear feasible but no reference was cited.

The GAEZ model has not specifically been developed for climate change impact assessments. Impacts on coffee production have not been analyzed. Figure 22 shows the distribution of agro-climatic yield potential in a high-input rain-fed system. Area with sufficient climatic conditions is shown to be drastically reduced by the 2050ies when comparing the baseline data with data from the Had-CM3-A2 GCM data.

Figure 22. GAEZ v3.0 *C. arabica* agro-climatic yield potential (High-input/Rain-fed)



Top: baseline; Bottom: 2050 conditions (Had-CM3-A2 scneario); dark green indicate high yield potential, yellow intermediate potential, red marginal zones; white areas have no yield potential (own illustration based on Fischer et al. 2012)[2].

Brazilian production is projected to migrate southwards, while area in Minas Gerais province appears to be lost. Yield potential is also reduced in the Congo basin, East Africa and Asia. Also Central America is negatively affected though it is difficult to evaluate the magnitude. Substantial area remains suitable in Ethiopia.

No model performance metric is provided for the GAEZ data. Assessing the data based on my own judgement shows an overall good fit of the current distribution on global level. However, some regions were underestimated, and others overestimated. Under current conditions e.g. production in Nicaragua, Brazil, Indonesia and the Philippines is more extensive than

[2] All map illustrations produced by myself use the geographic coordinate system WGS1984. The unit for each pixel or grid cell of the raster data is degrees of arc. One degree is equivalent to 60 minutes of arc. At global scale typically resolutions of 5arcmin or 2.5 arcmin were used. At sea level along the equator 1 arcmin is approximately 1.852km, each grid cell thus about 3.4km² or 342ha. However, geographical coordinate systems are true to their geographic unit but not true to distances or area. With increasing latitude the actual distance per arcmin is reduced. At 30° an arcmin equals about 1.600km, at 60° 0.900km. E.g. Europe therefore appears to have more area than it actually has in comparison to equatorial regions. But because coffee is grown at low latitudes between 22°N and 26°S, a few exceptions up to 26°N and 30°S, the geographic coordinate system is used throughout the thesis. Area in the extreme South of Brazil may therefore appear slightly overrepresented in map representations.

modeled by GAEZ. Also locations in Zimbabwe, Yemen, Jamaica, Nepal and West Africa were omitted. However, novel production areas in Southern China were described accurately. Yield potential in the Southern Congo basin appears overestimated. Currently no *C. arabica* production can be found there.

2.2.3.5 EcoCrop

Ecocrop is originally a database by the FAO that was based on expert opinion. It is openly available in the internet (FAO 2011). For about 2568 cultivated plants the database lists abiotic factors that determine crop performance. Implemented as a model the spatial suitability is assessed based on the following climate variables during defined growth season: Killing temperature, absolute minimum and maximum temperatures, optimum minimum and maximum temperatures and growing season precipitation found (Ramirez-Villegas, Jarvis and Läderach 2013).

The algorithm calculates suitability values based on two separate calculations, one for precipitation and the other for temperature. Beyond the above mentioned absolute thresholds, the suitability is given as zero, or unsuitable. Within optimum conditions suitability is a hundred percent. For areas where climatic conditions fulfill suboptimal requirements the model calculates suitability values on a range from 1 to 99. The suitability calculation is repeated 12 times, so that an optimal growing season may be found (Ramirez-Villegas, Jarvis, et al. 2013). For some applications parameters are not taken from the EcoCrop database. Data on the geographical distribution of crops is used to estimate parameters.

EcoCrop has been applied to model the impact of climate change on a large number of crops (Lane and Jarvis 2007). While the model is of great use in many applications, for coffee some of its assumptions are critical. For Arabica coffee the values for optimal growing season mean minimum and maximum temperature are estimated to be 14°C to 28°C, absolute minimum temperature is given at 10°C and max temperature 34°C. The respective precipitation values are given with 1400 to 2300mm/year and 750 to 4200mm/year (FAO 2011). Estimated values based on presence data are presented in Table 3.

Table 3. Ecocrop environmental limits for Arabica production

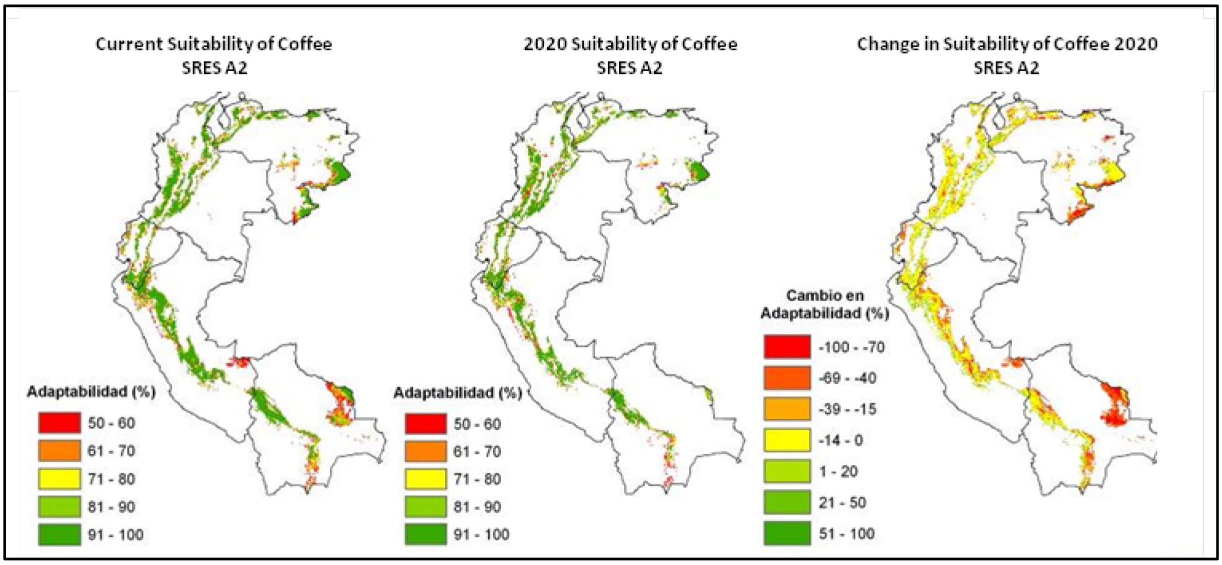
Temperature variables	**Killing temperature**	**Minimum temperature**	**Optimal min temperature**	**Optimal max temperature**	**Maximum temperature**
	0°C	16.5°C	18.6°C	22.7°C	24.7°C
Precipitation variables		**Minimum precipitation**	**Optimal min precipitation**	**Optimal max precipitation**	**Maximum precipitation**
		126mm	1154mm	3210mm	4238mm

(A. Quiroga, personal communication)

Ecocrop in its unadjusted version was used by Lane and Jarvis (2007) to model the global impact of climate change on the most important crops. Coffee ranked among the worst affected crops. Lane and Jarvis (2007) saw a decrease of worldwide suitable areas of -15% (HadCM3) or -7.3% (CCCMA) under the A2a scenario.

The Ecocrop crop model has also been used to model the impact of climate change along the Andes. As the map shows the suitability change until 2020 was generally negative (Figure 23).

Figure 23. Ecocrop suitability projection for Arabica coffee in South America



(Taken from Zapata-Caldas et al. 2011)

Especially lower altitudes were found to loose suitability, while higher altitudes may be better suitable for Arabica in the future.

2.2.3.6 Crop niche selection for tropical agriculture

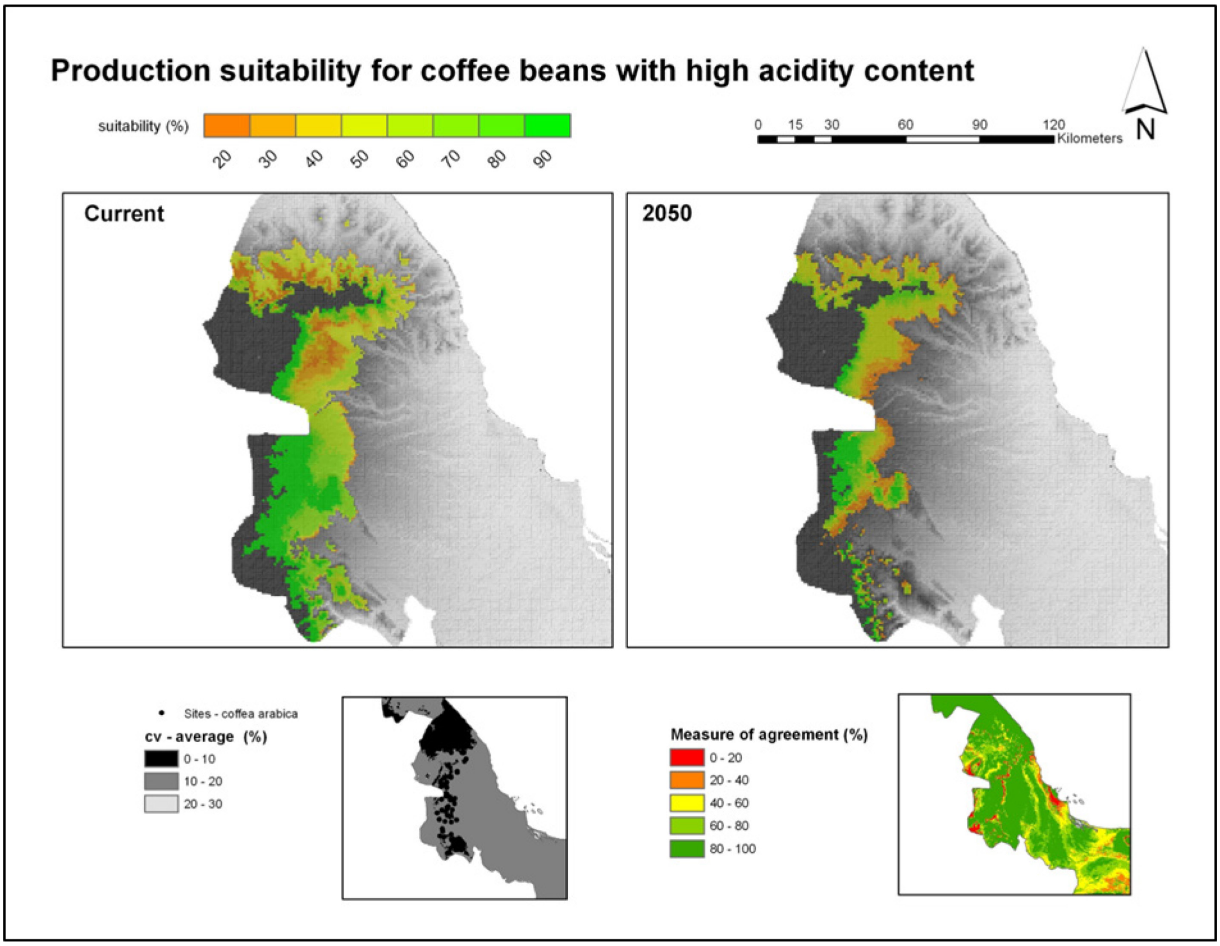
CaNaSTA (Spanish for basket) is a crop modeling surface based on Bayesian statistics and is an acronym that stands for Crop Niche Selection for Tropical Agriculture (Whitsed, Corner and Cook 2012). It incorporates both crop occurrence data and expert data. A key feature is the ability to predict performance.

In its Bayesian probability modeling approach CaNaSTA combines a prior and a conditional probability distribution to calculate a posterior probability. The prior probability is the probability that a crop is well suitable ("response variable"). This probability may be derived from occurrence data or expert opinion. CaNaSTA then derives a "joint probability" which is the probability of two events occurring together, such as crop suitability ("response variable") and a biophysical condition ("predictor variable", e.g. annual mean temperature). The posterior probability distribution finally is a function of the prior probability and the conditional probabilities.

The advantage of the model is the explicit incorporation and modeling of uncertainty. For the prior probability measures of uncertainty may be calculated or are stated by the experts. Two kinds of maps provide useful information: A map showing the most likely response value, and a map showing the certainty of this response.

CaNaSTA has been applied by CIAT for a wide set of applications in coffee such as occurrence of high quality of coffee or occurrence of coffee pests. In the case of quality distribution modeling the accurateness of the model has been confirmed by experts. The model predicted a region of high probability of occurrence of quality coffees different from the sites that were used for model calibration. The experts confirmed that the predicted site indeed is known for fine coffees (Oberthür et al. 2011). The model CaNaSTA was also used to project impacts of climate change on quality of coffee. Figure 24 shows an example of such an assessment from Veracruz in Mexico (Läderach, Oberthür, et al. 2011).

Figure 24. CaNaSTA projection of suitability for coffee beans with high acidity content



Large maps show the predicted suitability for high acidity content in Veracruz, Mexico, for current and 2050 conditions. The small maps show the coefficient of variation and the measure of agreement between predictions. High acidity content is used as a proxy to indicate potential for high quality coffee (taken from Läderach et al. (2011).

For Veracruz, Mexico, CaNaSTA projected an altitudinal migration of high quality coffee (Figure 24). Lower slopes that previously were climatically suitable for high acidity contents in coffee beans are projected to be unsuitable in the 2050 scenario. The model did not project novel areas in high altitudes. Rather, total area was reduced by the losses in low altitudes.

2.2.3.7 Maxent

The Maxent method (Phillips, Anderson and Schapire 2006) for species distribution modeling was used for several regional impacts on climate change effects on coffee. It will be used to develop a global climate change impact model in the following chapter. Therefore it is discussed with greater depth here.

Maxent uses the maximum entropy principle to predict the abiotic niche of a species, only using presence data (Phillips et al. 2006). Its output is an estimate of probability of presence (Phillips and Dudik 2008). It is generally agreed that its performance is better than other modeling methods when data is sparse and noisy (Elith et al. 2006). It has been criticized to be prone to biases in samples (Peterson, Pape and Eaton 2007) a problem that may occur in agricultural data. However, such biases may be dealt with (Anderson and Gonzalez 2011).

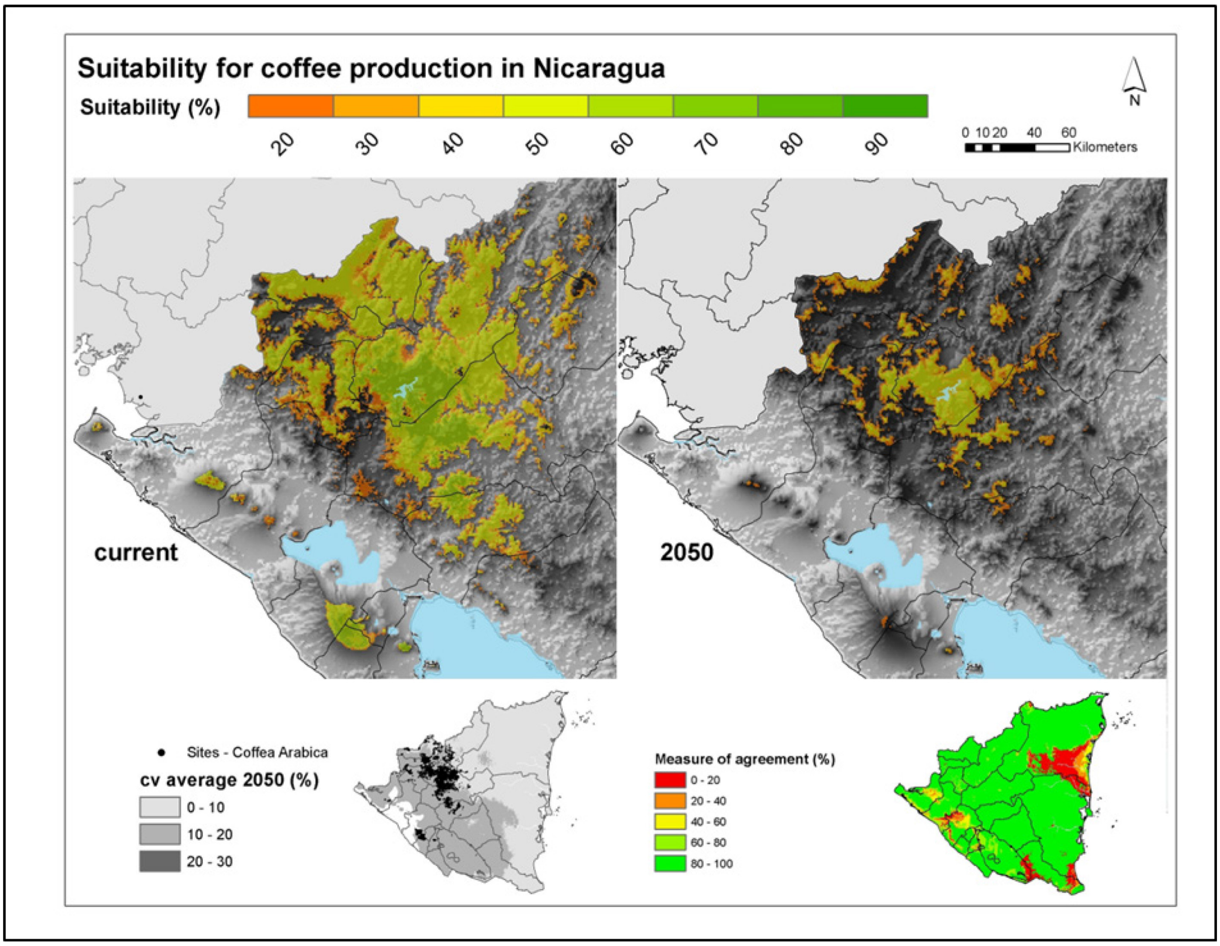
Input data for features may be categorical, discrete ordinal or continuous. Thus, the software may work with specifications such as e.g. soil type, soil quality or temperature. These features may be interpreted in a range of ways, for example as threshold values or linear relationships, attempting to realistically model the species distribution. Maxent has been found to be most accurate for species ecological niche modeling when data is limited (Hernandez et al. 2006; Elith et al. 2006; Ortega-Huerta and Peterson 2008). Result accuracy is further improved when geographical ranges are limited and environmental tolerance is low (Hernandez et al. 2006). However, Maxent also produces useful results when applied to cultivars with a wide ecological range and widespread use when the dataset is sufficiently large, even when the production system is highly socially and historically influenced such as rice production in Thailand (Heumann, Walsh and McDaniel 2011).

If applied to coffee niche distribution modeling Maxent reaches high predictive power. If presence data is specifically sampled for the project a careful sampling strategy may avoid biases, in which case the model has ideal prerequisites to predict the complete distribution. A common method to determine the performance of the model is to withhold a random set of presence data, typically 20%, and to test if the model predicts these presences correctly. A correct prediction rate higher than 50% is better than a random model, rates higher than 70% are usually considered a good performance in certain applications. With a good sampling strategy that is unbiased and a solid number of presence data Maxent reaches correct prediction rates of up to 95% of coffee presence data (Ovalle, Läderach and Bunn 2012).

Maxent very carefully extrapolates from the known presences to other areas without presences (Merow, Smith and Silander 2013). This may lead to the omission of parts of suitable areas. Especially if climate change creates zones with suitable climatic conditions that are remote from the original points of presence, Maxent may underestimate these. Together with the assumed niche conservatism (no adaptation to novel climatic conditions) Maxent therefore tends to overestimate the impact of global warming.

The Maxent method has been used in several impacts studies for coffee. The first application was to develop adaptation strategies in the Sierra Madre de Chiapas, Mexico, region (Schroth et al. 2009). For Nicaragua a strong decrease in suitability for Arabica coffee was projected, especially at low altitudes (Läderach, Lundy, et al. 2011) (Figure 25).

Figure 25. Maxent projection of suitability for coffee production in Nicaragua



The map shows the prediction of Maxent for current and future suitability for coffee production in Nicaragua. The small maps show the coefficient of variation and the measure of agreement between models for the model region. Black points indicate sample farms. Maps taken from Läderach et al. (2011).

A similar result was found for the entire Mesoamerica region (Läderach et al. 2009). Decreases in temperature seasonality and increases in precipitation seasonality were found to drive Arabica production to higher altitudes, with a generally more challenging climate in Kenya (Läderach et al. 2010). In Indonesia both regional migration and altitudinal migration may threaten forest covered land when Arabica coffee suitability in traditional regions is reduced (Schroth et al. 2014). Wild relatives and landraces of Coffea arabica are the focus of a study by (Davis et al. 2012). In this study a 90% reduction in area suitable for in situ conservation of coffee genetic resources was projected for the year 2080. A single study examined the impacts of climate change on Robusta coffee in Vietnam. A serious reduction in area suitable, and an altitudinal migration was reported (Läderach et al. 2012).

2.2.4 Summary of modeling approaches

In the literature several studies could be found that attempted to assess the impacts of climate change on coffee production (Table 4). Nearly all models were confined to a limited regional scale and cannot be readily compared because of their differing model parameters. A global model, the GAEZ model, was found to be unreproducible. Therefore, here the previous modeling approaches are compared with respect to their usefulness for a global application.

Table 4. Summary of previous research on the impacts of climate change on coffee

Approach	Region	Result	Species	References
Regression	Veracruz, Mexico	Yield reduction/abandonment	*C. arabica*	Gay Garcia et al. 2006
	Tanzania	Yield reduction	*C. arabica*	Craparo et al. 2015
Simulation	Costa Rica	Yield reduction	*C. arabica*	Van Oijen et al. 2010b
Envelope	Uganda	Altitudinal migration	*C. canephora*	Simonett 1988
	Ethiopia	Negative suitability changes	*C. arabica*	Rüegsegger 2008
	Brazil	Latitudinal migration/Regional migration	*C. arabica*	Assad et al. 2004; Zullo Jr et al. 2007; Assad and Pinto 2008; Zullo Jr et al. 2008; Zullo et al. 2011
GAEZ	Global	Reduced area/latitudinal migration	*C. arabica/ C. canephora*	Fischer et al. 2012
EcoCrop	South America	Altitudinal migration	*C. arabica*	Zapata-Caldas et al. 2011
	Global	Severe impacts, -17% area	*C. arabica*	Lane and Jarvis 2007
CaNaSTA	Veracruz, Mexico	Reduced quality/ Altitudinal migration	*C. arabica*	Läderach et al. 2011
Maxent	Nicaragua	Reduction in suitability/ Altitudinal migration	*C. arabica*	Schroth et al. 2009
	Mesoam.	Altitudinal migration	*C. arabica*	Läderach et al. 2009
	Kenya	Altitudinal migration	*C. arabica*	Läderach et al. 2010
	Ethiopia	90% loss of area	*C. arabica*	Davis et al. 2012
	Vietnam	Reduction of area/Altitudinal migration	*C. canephora*	Läderach et al. 2012
	Indonesia	Regional migration/ Altitudinal migration	*C. arabica*	Schroth et al. 2014

(Own data)

Models based on intra-annual variation have the capability to distinguish differential effects of temperature and precipitation variation during several stages of growth. In its most elaborate case, the dynamic process model Caf2007, the approach is able to include even management practices that allow the explicit modeling of adaptation practices. However, such models are limited by the lack of globally consistent high-resolution data of both climate and yield. A global application of the correlation approach with a set of local models would require substantial efforts to gather and prepare the necessary data. Such a project would likely be highly rewarding in terms of knowledge gained. Nevertheless, the extrapolation of the resulting models in larger time scales could still be problematic. On the other hand, the process model Caf2007 would allow such application to future climate data. However, this model has only been calibrated for the specifications of production systems of the Turrialba region in Costa Rica. Applying this model to other regions and different growing conditions would need substantial research.

The spatial models are more diverse in their approaches and complexity. Some models were found to be rather simple, e.g. a combination of two or three variables with expert defined limits. The GAEZ model extends this approach to high complexity with several variables. The objective of this approach therefore went beyond the estimation of the spatial distribution of suitability indicators for the crops modeled: the suitability indicator was translated into yield potential by applying penalties to a maximum achievable yield if not all variables fulfilled optimality conditions. The resulting model appeared to estimate well major production regions, but underestimated minor ones. Because of a lack of documentation this approach cannot be reliably evaluated and applied to different climate change scenarios.

Another group of approaches therefore attempted to calibrate the environmental limits of coffee production based on observed occurrence locations. E.g. (Rüegsegger 2008) and the Ecocrop model used this approach. The resulting models of the spatial distribution of the suitability for coffee production were crop specific maps of an unweighted combination of a few variables with often limited relevance to crop production. Therefore the Bayesian and machine learning approaches CaNaSta and Maxent appear to be advantageous because of their ability to handle a large number of variables. CaNaSta is the less researched algorithm of the two. Maxent has been applied in a high number of cases and a large number of publications can be accessed (Merow et al. 2013). CaNaSta furthermore requires spatially explicit production data to estimate the distribution of coffee which was found to be a disadvantage. With much more reduced data requirements Maxent has been shown to reliably model the spatial distribution of species (Elith et al. 2006).

These properties of previous modeling approaches for coffee and climate change are summarized in Table 5. An ideal modeling approach would have low data requirements, but could make use of a high number of variables and fit complex functions to estimate yield in a highly reproducible way. Regression or simulation approaches arguably have prohibitively high data requirements for a global scale model. Envelope and EcoCrop models were found to be of very low complexity and handle only few variables so that their ability to reflect the crop-climate relationship is limited. The GAEZ approach was found to be unreproducible because of the lack of documentation. The use of expert knowledge and its static software implementation makes the reproduction of CaNaSTA results difficult. The machine learning algorithm Maxent will therefore be used to model the biophysical impacts of climate change on coffee production. The data requirements are feasible for a global study, it can handle several variables, fits models with high complexity and results are easy to reproduce. However, yield itself is not estimated but only probability scores indicating the whether a pixel is suitable for coffee or not.

Table 5. Summary of the properties of previous modeling approaches in coffee and climate change

Approach	Data requirement	Number of variables used	Complexity of functions	Yield estimation	Ease of reproduction
Regression	Very high	Few	Intermediate	Yes	Difficult
Simulation	Extremely high	Very high	Very High	Yes	Difficult
Envelope	Low	Low	Low	No	Easy
GAEZ	High	High	Intermediate	Yes	Difficult
EcoCrop	Low	Low	Low	No	Easy
CaNaSTA	Intermediate	Intermediate	Intermediate	Yes	Difficult
Maxent	Intermediate	High	High	No	Easy

(Own data)

Independent of the approach all models found net reductions in area available for coffee production, or projected reductions in yields (Table 4). Data is not quantitatively comparable because different time horizons and emission scenarios were used. However, a number of trends can be identified: regional abandonment of coffee production, migration in latitude, and migration in altitude.

The studies that suggested a complete abandonment of coffee production were both of the regression type (Gay Garcia et al. 2006), and the Maxent type (Davis et al. 2012), and looked at different time horizons (2020 vs. 2080). Latitudinal migration was suggested by the envelope type approaches that were based on annual mean temperature ranges on global scale, in the Brazilian South and the Ethiopian North (Lane and Jarvis 2007; Rüegsegger 2008; Zullo et al. 2011; Fischer et al. 2012). The majority of studies however, concluded that coffee production will migrate in elevation, mostly due to losses at lower elevations and some gains at higher elevations (Simonett 1988; Schroth et al. 2009; Läderach et al. 2010; Läderach, Lundy, et al. 2011; Läderach et al. 2012; Schroth et al. 2014).

Because of the regional nature of the studies little and the differences in modeling approaches it cannot be concluded whether the impacts differ between regions. Furthermore, studies were inconclusive in their comparison of impacts on the two coffee crops Arabica and Robusta. And, little attempt was made to link impacts to economic indicators, e.g. viable yields or actual area under production.

2.2.5 Conclusion

This section established the assumption that global coffee production will be affected by climatic changes. Both coffee species that dominate production of coffee require specific climatic conditions in order to be productive. The Arabica species is heat sensitive and Robusta sensitive to cold temperatures. Several studies have investigated the impacts of climate change on regional scale. Despite methodological differences the projected results have always been negative: latitudinal expansion, lost area at low altitudes to complete abandonment of coffee production.

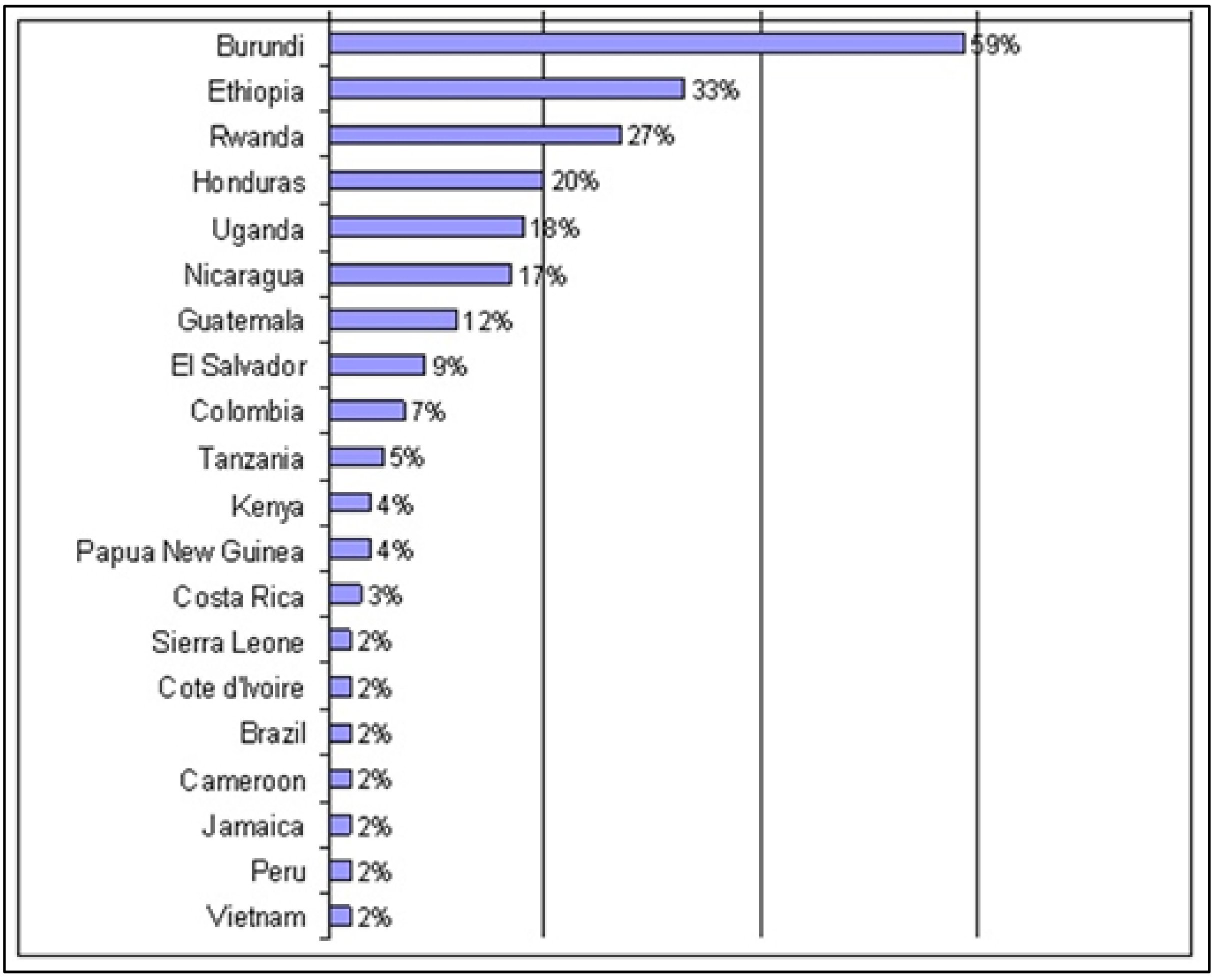
Nevertheless, most studies were found to be limited to regional scales and did not estimate economic impacts. Therefore, results were not comparable between regions and species and did not include economic feedbacks. It was therefore proposed to use the machine learning classification algorithm Maxent to develop a global model of the climate change impacts on coffee production before integrating the results with the partial equilibrium model Globiom.

2.3 Coffee consumption and production

Coffee production and consumption is very labor intensive. Beginning with largely manual sowing, cultivation, and harvesting to fancy coffee shops with hand brewed specialty coffee some estimate that the entire coffee supply chain provides a livelihood for 125 million people worldwide (Pendergrast 2010). Only on the production side more than 26 million workers are employed in low and middle income countries in coffee production (ICO 2010). Assuming a family size of four in the producer countries a 100million people could be dependent on this crop.

In several countries coffee generates more than 10% of export earnings (ICO 2014). The indirect dependence of rural societies on coffee production could be even larger than the direct employment. In Ethiopia more than 700,000 families are involved in coffee production and a total of up to 15million people depend on this sector directly or indirectly in this economy (Vega, Rosenquist and Collins 2003). According to estimates by the International Coffee Organization (ICO) there are currently (2000-2010) 7 countries that make more than 10% of their export earnings from coffee export (Figure 26). This number is half of what it was in the ’96-’00 period but coffee nevertheless remains a significant source of tax income and foreign exchange earnings in these countries (ICO 2014).

Figure 26. Average share of coffee exports in total export earnings (2000-2010)



(taken from ICO (2014)

The regional importance of coffee production is exemplified by the impacts of the “coffee crisis”- a decade of historically low coffee prices on commodity exchanges caused by an unfortunate combination of factors (Varangis 2003). About 28% of the rural labor force in Central America derives employment from the coffee sector. In connection to the crisis seasonal employment in coffee production in Central America declined by 20% and permanent employment by 50%. A total of 540,000 workers lost their primary income during the crisis. Many of these jobs were lost in rural areas without income alternatives so that this also impacted health care and education in affected regions (Varangis 2003). The crisis furthermore compromised the region's ability to cope with future crisis as also funding for research depends on the revenue that comes from sales. E.g. at Colombia's CeniCafe coffee research institute the staff was reduced from 436 to 169 during the crisis (Vega et al. 2003).

Coffee as a luxury beverage has been called "the second most traded commodity after oil" (Pendergrast 2010). This title is questionable (Pendergrast 2009) but reflects well its property that most of the produced coffee is not consumed at origin, but in urban service societies. Demand and production have been increasing over the past decades. Lately especially the increasing Robusta production has changed the field substantially. Here the trends on coffee markets are briefly reviewed.

2.3.1 Historical development of the market for coffee

A little known fact is that coffee is the eponym of an entire family within the plant kingdom. The coffee family *Rubiaceae* comprises an estimated 13.000 species. *Coffea arabica* is only the most prominent in this family because of its importance for commercial production. The name "Arabic coffee" describes the historical trade route through which it became known to Occidental Europeans in the late 17th century. Consumption in the near East can be traced back to the early 15th century. Here, the dark brew received its name from the Arabic word *qahwah,* which is usually interpreted as a type of "wine" (Crawford 1852). Other sources however see the word to refer to the region of origin *of Coffea arabica*, the kingdom of Kaffa, today's Ethiopia.

The very high prices of the new beverage and the rise of colonialism lead to the introduction of coffee seed by colonialists in tropical regions outside of its region of origin. In Ethiopia they found what the locals called "bunn" or "bunno" (Bleke in (Crawford 1852)[3]: The coffee shrub and its seed. This germplasm was transported to the colonies in an effort to profit from coffee trade. No early source on coffee fails to mention the importance of a suitable climate

[3] Today's referral to "bean coffee" stems back to an effort to differentiate from substitute coffee products. The expression "bean" borrows from the Amharic "bunn".

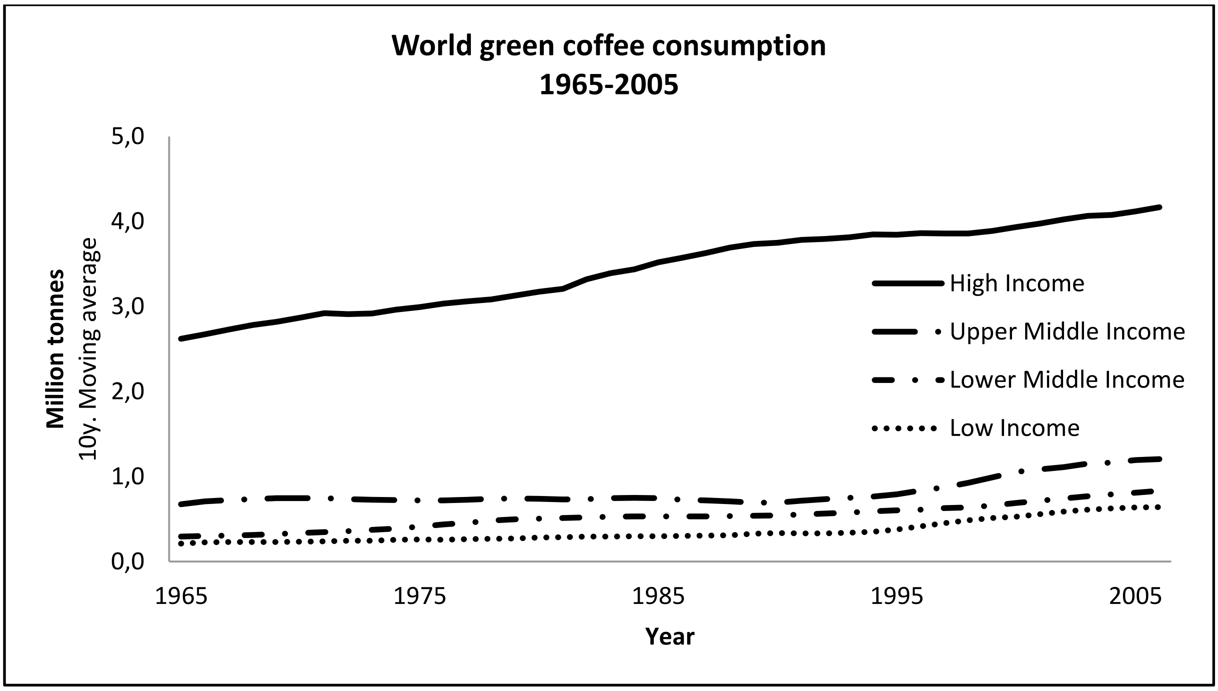
for the successful introduction of plantations. The second half of the 18th century saw quick expansions of plantations in Java and Brazil where colonialists found ideal conditions. In an effort to explain the success of coffee as compared to other luxury beverages (Crawford 1852) emphasizes that in the colonies “even the Indians of Sumatra” can produce good quality coffee without the help of Europeans.

By the mid-19th century Brazil was already the world’s largest producer at an estimated 36% of world production, followed by Java at 25% (Crawford 1852). At the beginning of the 20th century this picture had changed with the colonialization of the vast areas of Minas Gerais state in Brazil that are suitable for coffee production. In 1913 about three quarters of the world’s coffee crop were grown in Brazil (Anonymous 1913).

2.3.2 Current demand

The average world consumption of coffee in the ’01-’10 decade has been about 6.7 million ton a year. This is equivalent to a bit over 1kg per person and year (FAO 2014b). The total consumption of green coffee has grown steadily over the past decades. Historically, more than two thirds of all coffee was consumed in North America or countries of the European Union. Even though this gap appears to close because upper middle income countries are increasing consumption, the split of coffee supply across countries remains very uneven. The 15% of the global population that inhabits high income countries consumes about 60% of global coffee (Figure 27).

Figure 27. Development of world green coffee consumption by income group ('65-'05)



Data shows million ton/ year (10y moving average) by income group[4] (FAO 2014b).

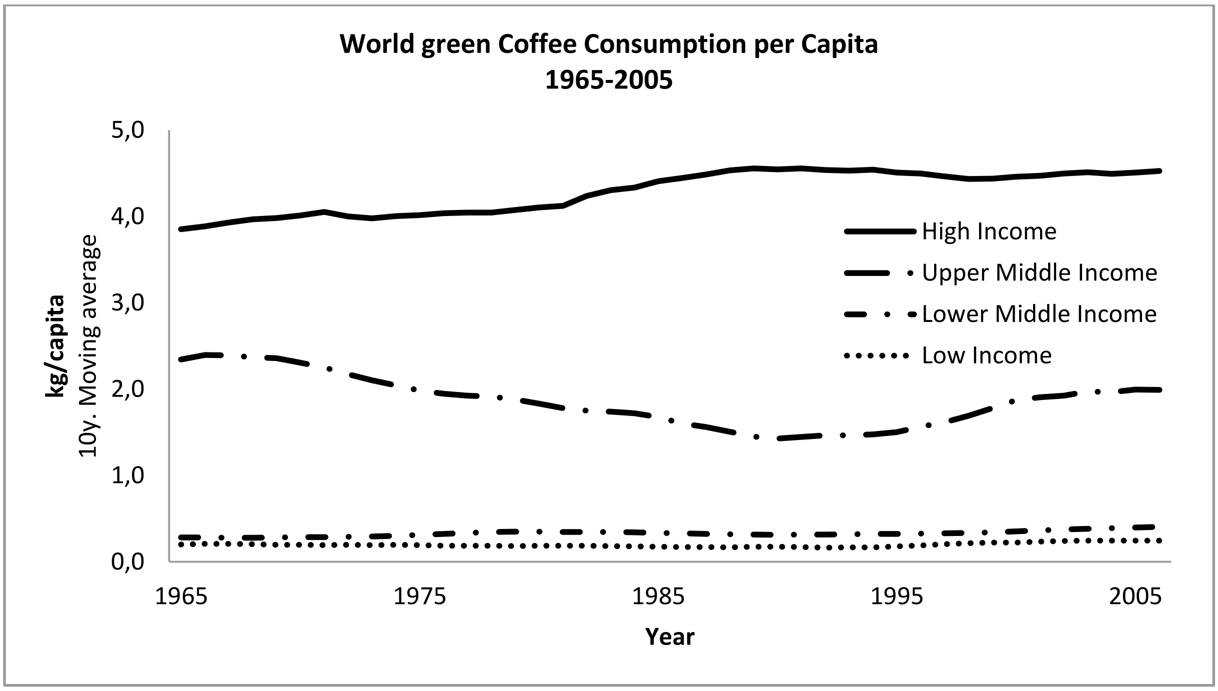
Despite this increase in total demand of about 2% per year over the past decades coffee markets are often described as stagnant. This is because there has been little increase in per capita consumption over the past two or three decades in high income countries (Figure 28). These traditional coffee markets consume a stable 4.5 kg/person, with some markets such as the United States drinking up to 6.5kg/person/year. Per capita consumption in low income and lower middle income countries also remained stable but at comparatively low levels. Here, increasing total demand possibly resulted only from population growth. In many countries little to no coffee is consumed (Figure 28). Interestingly, per capita consumption in upper middle income countries has declined over several decades but has more recently increased again. This group of countries also comprises several Latin American countries, some of which are major producers like Brazil, the world's most important coffee source.

The two dominant trends, stagnating consumption in high income countries and increasing consumption in upper middle income countries, have been met with increasing product differentiation from the side of processors (Lewin, Giovannucci and Varangis 2004). At both ends of the quality range new products are being developed. On high income countries coffee products are increasingly certified for Fair Trade or for ecological production practices. On the other end the first products to enter new markets are often low-quality convenience

[4] Income groups were based on income data from FAO (2014) and income group definitions from Soubbotina (2004)

products. Consumers in emerging markets preferably buy soluble coffees that are based on low-quality *C. canephora var. Robusta* green coffee.

Figure 28. Development of world green coffee consumption by income group (1965-2005)



Data is in kg/capita/year (10y moving average) by income group[4] (FAO 2014b)

Considerable uncertainty prevails about the future development of the global coffee demand. Coffee is generally seen to be popular in urban societies with high GDP's. The recent increases in demand in upper middle income countries are seen by some as the start to jumps in demand as these economies are rapidly urbanizing and adding GDP. However, others warn that this might be a short term bump caused by historically low prices due to overproduction. Nevertheless, the low per capita consumption in many middle income countries, the increasing urbanization and global growth of GDP are interpreted as a large market potential for future growth by the industry (Lewin et al. 2004). With its large population, little per capita consumption but quickly growing economy the Chinese market is the greatest source of uncertainty. On both ends of the quality range the market appears to be growing. Soluble coffees are increasingly popular, and also American coffee house chains are expanding business. Nevertheless, per head consumption was still tiny at 30gr/person in 2009 (about 4 cups of coffee) (International Trade Centre 2011).

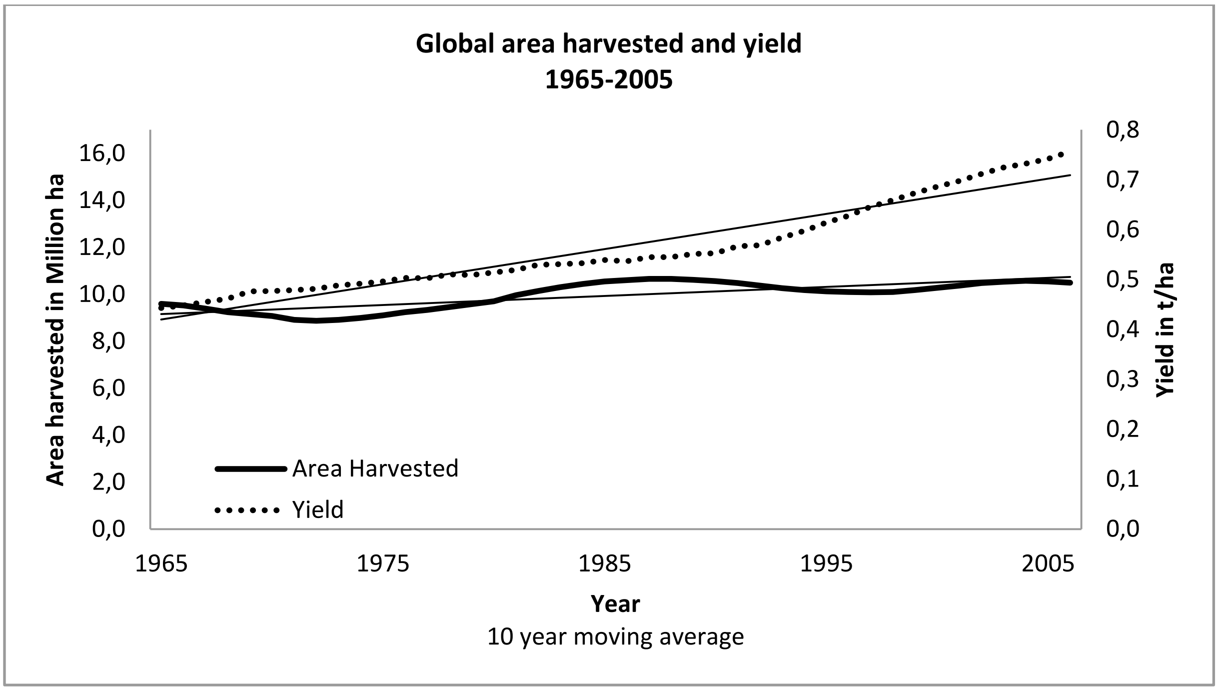
The overall market is thus growing. Over the past decades this growth came not as much from an increase in per capita consumption but rather from more consumers. The past decade has seen some changes to this pattern as upper middle income countries have seen higher per capita demand. It is not sure however, to which extent this trend continues. Furthermore, most of this growth has gone into low-quality Robusta based soluble coffees.

2.3.3 Current production

Nearly all coffee production is based on two species: *C. arabica* ("Arabica coffee") and *C. canephora* var. Robusta ("Robusta coffee"). About 10 million hectares are planted with these two species. Today Brazil still ranks as the world's largest producer of green coffee, although its dominant position has been challenged recently by Vietnam. While the former is a major producer of Arabica coffee, the latter produces predominantly Robusta.

A shown above, consumption has about doubled over the past decades (Figure 27). This demand has been met by producers with increased production, sometimes over-production (Vega et al. 2003). Most of this increase in output has not come from additional area but from an increase in productivity (Figure 29). The overall market has grown by about 2% per year. Approximately 0.4% of this growth has come from additional areas and 1.6% from yield increases.

Figure 29. Development of global coffee area harvested and yield (1965-2005)

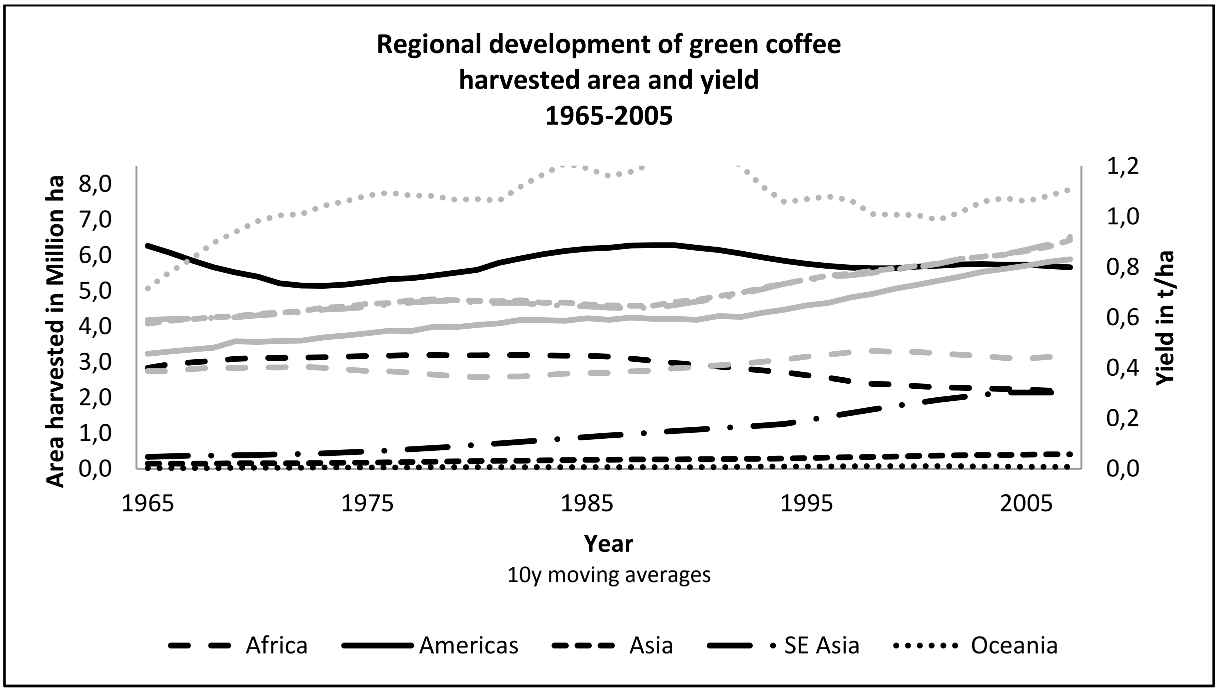


Area in million ha and yield in t/ha (10y. moving averages); narrow lines indicate the trend over the '65-'05 period (FAO 2014b).

A regional split of the same data shows that not all regions were equally able to profit from the additional demand (Figure 30). Most regions improved yields per hectare with the exception of Africa. The other two major coffee regions, America and Asia have seen a near

100% increase in average yields per hectare, while in Africa yields remained at 1960ies levels. Possibly as a consequence of this harvested area in African countries was reduced by about a third over the past decades. At the same time area in the Americas has remained roughly the same area and increased substantially in South East Asia (Figure 30).

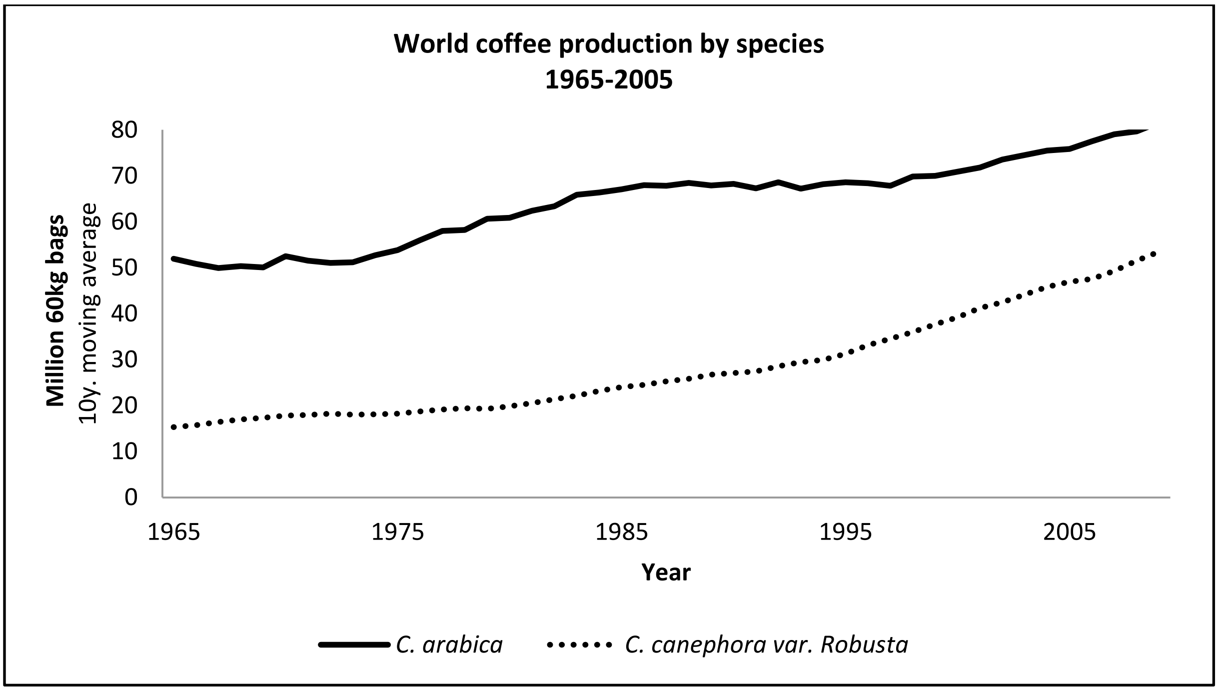
Figure 30. Development of coffee area harvested and yield by region (1965-2005)



Area in million ha and yield in t/ha (10y. moving averages); black lines represent area harvested, grey lines yield per ha (FAO 2014b).

Part of this development was a disproportional growth of Robusta output. Generally, this species produces a lower quality but a higher yield. Especially on emerging markets coffee of this species finds consumers as soluble coffee. When before 1970 less than 25% of global coffee was of the Robusta kind, it is now about 40% (Figure 31).

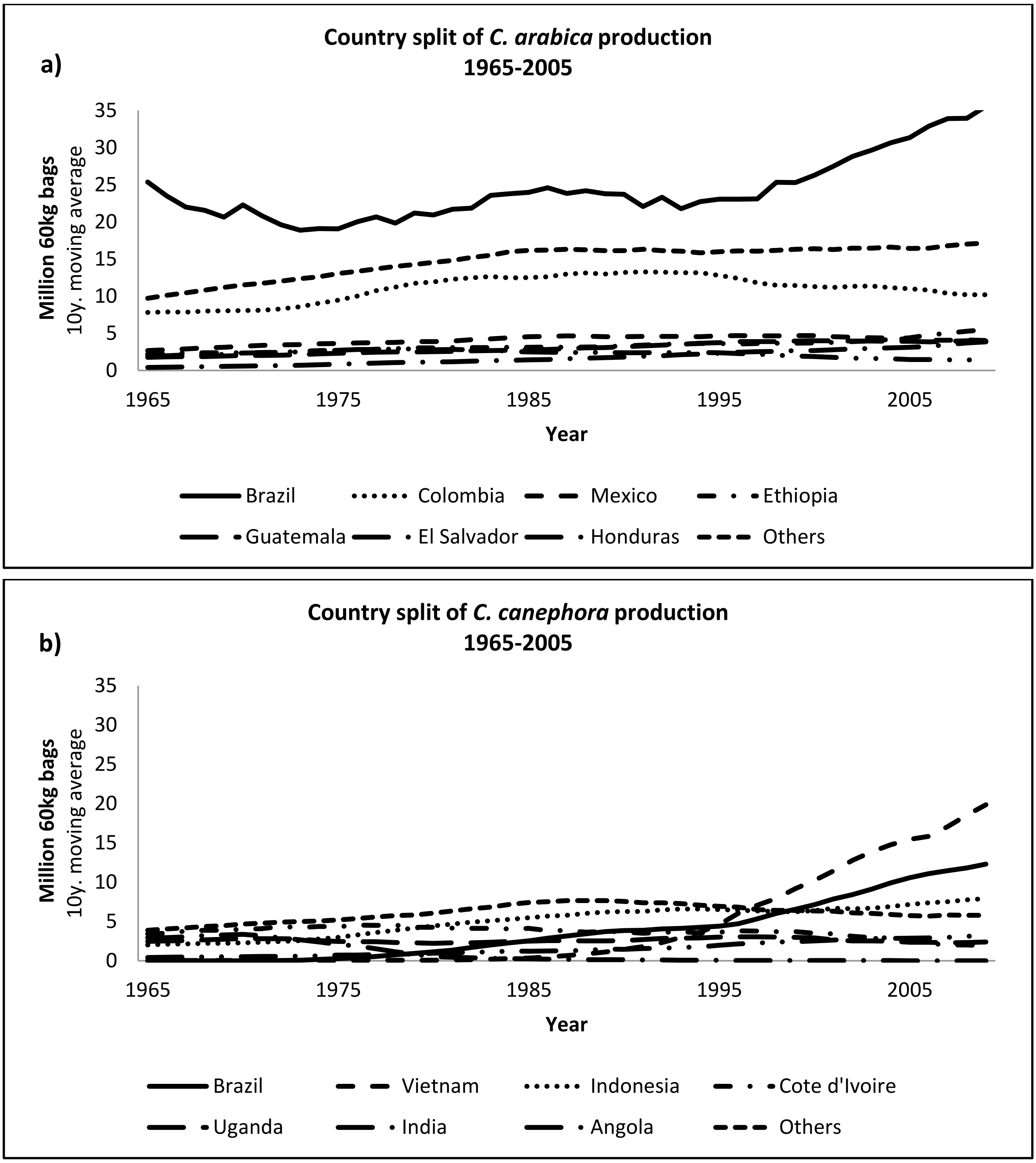
Figure 31.   World coffee production by species (1965-2005)



Million 60kg bags (USDA 2012).

A more detailed look at the major producer countries shows that nearly all of the increase in production has come from two countries: Brazil for Arabica and Robusta; Vietnam for Robusta. Figure 32 shows the development of produced quantities over time for the 7 largest countries and "others" for each coffee species. A sharp increase in Arabica production can be seen for Brazil after the 1990ies (Figure 32a). Other countries have not expanded production in a similar manner. This behavior is only matched by Vietnam's increase in Robusta production which increased from nearly no production to the second largest output worldwide (Figure 32b). On the Robusta market Vietnam and Brazil significantly increased production while other major producer countries have seen declines or unchanged outputs.

Figure 32. Largest producer countries by species (1965-2005)



Million 60kg bags a) *C. arabica*; b) *C. canephora* (USDA 2012).

Thus, the increase in total demand was matched by producers by increases in productivity. Global yields grew by nearly as much as demand so that area remained constant. African producers however did not increase productivity and hence coffee area has been reduced on this continent. The largest share of production increases was matched by only two countries: Brazil and Vietnam. While Brazil also increased Arabica production, Vietnam makes up for the majority of increases in low-quality Robusta production.

2.3.4 Market dynamics

As pointed out above, coffee can be distinguished into several qualities. On saturated markets retailers increasingly attempt to raise demand by product differentiation (Lewin et al. 2004). On the upper end bidders compete for single bags of high quality specialty coffee. However, nearly all coffee is traded on commodity stock exchanges (most notably London and New York). On stock markets coffee qualities are distinguished according to ICO specifications. The lowest price receive “Robustas”, all “Arabicas” receive higher prices. For Arabica the qualities “Colombian milds”, “other milds” and “Brazilian naturals” are distinguished. The Colombian milds receive the highest prices, Brazilian naturals the lowest (ICO 2013). The final product, ground coffee sold at supermarkets, commonly contains a mixture of various qualities, e.g. Robusta as a filler, Brazilian naturals for body, and Colombian milds to improve taste. The coffee market, even for commodity qualities, is thus differentiated by prices and qualities.

However, for it to be defined as a single market a price change in one of the qualities must induce similar price changes in the other qualities, so that a common stochastic trend for all qualities results. (Ghoshray 2010) demonstrated that prices do co-integrate when allowing for non-linear autoregressive effects. He argued that despite the imperfect substitutability and transaction costs the arbitrage effect increases with increasing price differences, but effects may be lagged. The different coffee qualities thus do form a single market.

The effects of the price increases and decreases on supply and demand were subject to a number of studies following the end of quota regularizations in the international coffee agreement (ICA) after 1989. For a period of three decades this agreement had kept prices at relatively high levels. After its end, coffee prices collapsed by approximately 40%. In their analysis of the impacts of this market shock on producer countries (Akiyama and Varangis 1990) report extensively on the short- and long term elasticities of supply, and the price elasticities of demand, and income elasticity of demand.

In line with other values found in the literature they reported low (0.04) short term price elasticities of supply after one year. With increasing duration after price shocks supply elasticity was found to increase. The ten-year price elasticity of supply was found to be in the range of 0.14 to 0.36 for most countries, and between 0.75 to 0.95 for some exceptions (e.g. Colombia and Ivory Coast) (Akiyama and Varangis 1990). These values are similar to what other sources report. E.g. Lewin et al. (2004) report a long-term elasticity of Latin American

Arabica production of 0.15 and 0.2 Robusta markets. (Otim and Ngategize 1993) report values between 0.052 and 0.526 respectively for Ugandan Robusta and Arabica production.

The same study by (Akiyama and Varangis 1990) reported a price elasticity of demand of about -0.2 on world markets, and individual price elasticities between -0.08 and -0.54 for European countries and the USA. In producer countries the price elasticity of demand was found to be lower at around -0.1 (Akiyama and Varangis 1990). Lewin et al. (2004) reported a somewhat lower value of -0.1 for the USA.

The income elasticity of demand for the period 1968 to 1986 was found to range from 0.1 to 2.89 for importing countries, and between 0.18 and 0.5 for exporting countries. The global average income elasticity of demand was 0.6.

2.3.5 Conclusion

Coffee production is highly regionally aggregated. Only a few major producing countries provide nearly all of the coffee consumed. Within these countries coffee production is confined to a few regions based on the climatic conditions. Even though the crop occupies limited area it is of high importance in these regions. For some countries it provides large shares of export revenues, or the crop has become an integral part of landscapes and culture.

Demand has grown steadily over the past decades. This was driven by population growth and increases in income. The projected increases in population and GDP as assumed by the SRES and RCP scenarios will therefore likely also affect coffee demand. In the past, changes in comparative advantages have shifted the regional distribution of coffee production. Climate change will again shift these relative advantages so that the regional composition of coffee origins will likely change as a result.

2.4 The integrated modeling framework Globiom

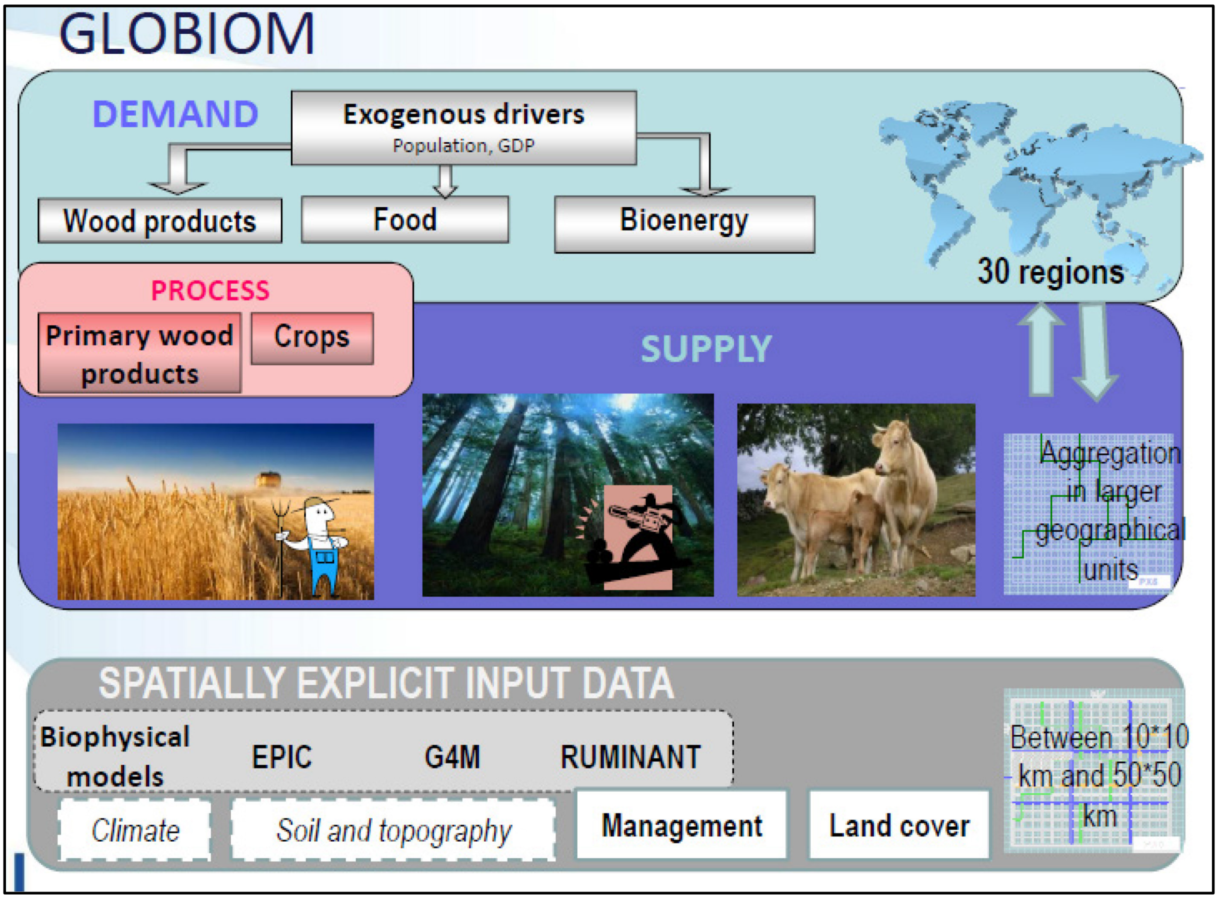
Previous sections in this chapter have shown that climate change will not only change the comparative advantages of regions for coffee production. The assumed changes in population and global income will also change demand. An analysis of the climate change impacts on global coffee production would therefore be incomplete without taking into consideration market effects. These effects should be modelled in an integrated fashion using a partial equilibrium model of the agricultural sector. Here, the Globiom integrated modeling framework will be introduced and the rationale for choosing this model is presented. Also, data needs to include coffee in this framework will be demonstrated[5].

In partial equilibrium models only a part of the economy is considered, in this case agriculture. Models assume rational behavior: producers maximize their profits and consumers their utility and demand and supply are in equilibrium for the sector that is regarded. Computable equilibrium models (CGE) also consider other sectors of the economy. They therefore allow feedback effects between sectors. In an agricultural context factor markets are often of interest. Food security impacts of changes in the equilibrium affect different sections of the population in diametrically differing ways depending on the household structure. However, CGE models require the formulation of equations for each unit of simulation. This often limits global CGE models to country scale. In a climate change context this decisively limits the insights that can be obtained as it is difficult to aggregate the impacts of climate change in a meaningful way on country level (Dumollard, Havlík and Herrero 2013).

On the other hand partial equilibrium models commonly rely on optimization approaches. An objective function that maximizes e.g. global welfare is maximized subject to a set of constraints. Constraints can be formulated on various levels so that detailed disaggregation is feasible. Therefore, here a spatially explicit PE model of the agricultural sector will be used. In comparison with other models of the same kind Globiom has several advantages: Globiom is global, spatially explicit with high disaggregation, non-monetary flows are included, and in addition to crop production also the livestock sector is modeled and a forestry model is attached (Figure 33).

[5] A formal description with definitions of variables, functions, parameters and indices can be found in (Havlík et al. 2011). Here, only aspects immediately relevant for the inclusion of coffee in Globiom are discussed.

Figure 33.   Schematic overview of the Globiom model



Taken from Mosnier et al. (2012)

2.4.1   Globiom model structure

Globiom is a recursive dynamic partial equilibrium model of the agricultural sector that includes crop, livestock, forestry and bioenergy production (Havlík et al. 2011). Globiom thus allows to model the effects of demand side changes and supply side technical changes on the spatial distribution of crop production. The model is recursive dynamic because scenario simulations are solved in 10 year steps with changes in one period being transferred to following periods.

Global welfare is maximized by optimizing land use and processing activities subject to resource, technology and policy constraints. Market equilibrium is calculated using constant elasticity demand functions and geographically explicit Leontief production functions to model supply. Globiom models commodity markets for 18 crops, livestock product calories, 6 different forest commodities and 7 bioenergy types (Schneider et al. 2011) (Table 6).

Table 6. Crops, livestock, forestry and biofuel products in Globiom

Group	Commodities
Crops	Barley, cassava, chickpeas, corn, cotton, dry beans, ground nuts, millet, oil palm fruit, potatoes, rapeseed, rice, soya, sorghum, sugarcane, sunflower, sweet potatoes, wheat
Livestock products	Animal food calories with fixed proportions of bovine meat, pig meat, sheep and goat meat, chicken meat, equine meat, fresh milk, turkey meat, and eggs from hens and other birds
Forest commodities	Primary products: saw logs, pulp logs, fuel wood, other industrial logs, processed products: sawn wood, wood pulp
Other commodities	Biodiesel, ethanol (1st and 2nd generation), methanol, heat, power, biogas

(Schneider et al. 2011)

Before solving, Globiom is calibrated by adjusting the cost parameter. Solving the model in an uncalibrated state leads to deviations of model solution from observed values where input data is inaccurate, inconsistent, or not obtainable in sufficient detail at global scale, e.g. crop rotations, machine costs etc. Additionally some important economic variables are not captured by the model such as market imperfections or quality differentiation. Microeconomic theory states that marginal revenues equal marginal costs. Therefore the model can be calibrated before solving by adjusting the costs in each country and management system to baseyear commodity prices. This brings the baseline solution closer to observed values.

2.4.2 Demand in Globiom

On the demand side input data is provided for 30 international regions (for region definition see Annex 1. Region definitions in Globiom). The region definitions are quite flexible and can be adapted to specific research questions.

For each region demand prices and quantities are taken from FAOSTAT. Between these regions net trade is accounted for. Further model external information defined for each region is population, GDP, dietary patterns, bioenergy demand, processing costs and coefficients and trade costs. To model future demand scenario assumptions about GDP, population and the development of demand have to be provided to Globiom. In previous work e.g. different development pathways for meat or bioenergy demand were compared (Havlík et al. 2011; Havlik et al. 2012). Also the effects of consumer policies to mitigate the impacts of climate change on food security have been studied (Mosnier et al. 2014). To model coffee such demand scenarios will be developed in chapter 5 from the SRES population and GDP projections.

2.4.3 Supply in Globiom

Supply side data is provided in much more detail. Spatially explicit input data in Globiom is organized based on the concepts homogenous response units (HRU) and simulation units (SimU) developed by (Skalsky et al. 2008). Input data for Globiom was compiled from a multitude of sources with differing spatial resolution and quality and therefore had to be harmonized. On HRU level parameters that are invariant over time were combined. Combination of HRUs with administrative information, land use data and climate information results in SimUs.

A HRU is constructed by combining a 5’ geographical reference grid (approx. 10km at the equator) with coarse data of landscape characteristics for slope, altitude and soil class. Five different altitude classes and seven slope classes are discerned. For soils, 6 different classes are used that describe the soil characteristics (Table 7). The result were zones with the same altitude, slope and soil class with 5’resolution.

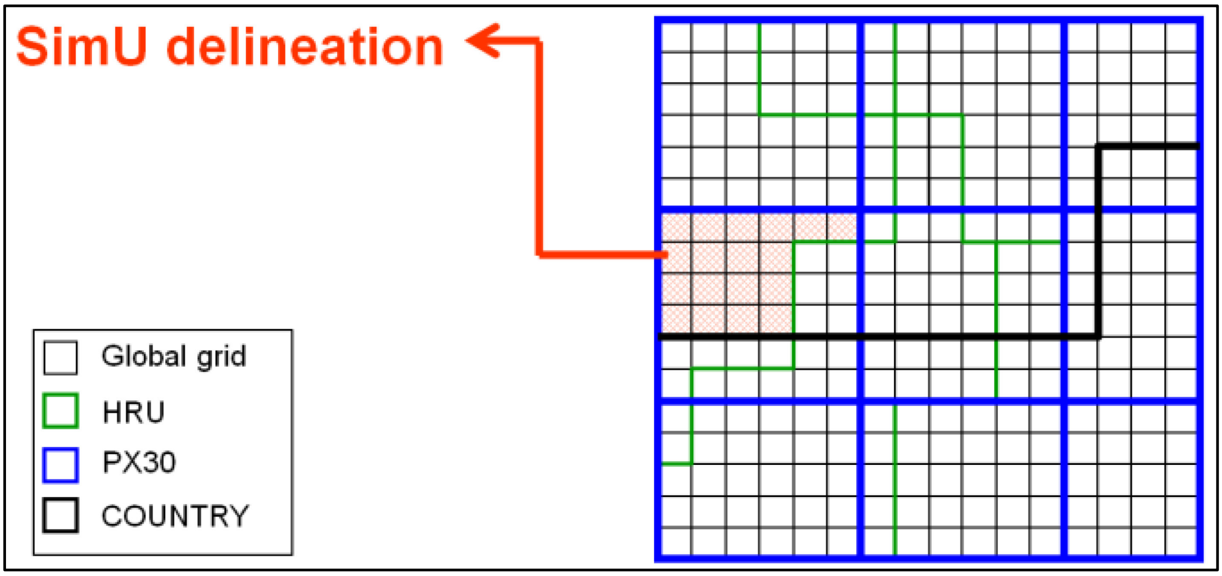
Table 7. HRU characteristics

Land characteristic	Unit	Classes						
Altitude	Meter	**1**	**2**	**3**	**4**	**5**		
		0-300	300-600	600-1100	1100-2500	>2500		
Slope	Degree	**1**	**2**	**3**	**4**	**5**	**6**	**7**
		0-3	2-6	6-10	10-15	15-30	30-50	>50
Soil		**1**	**2**	**3**	**4**	**5**	**6**	
		Sandy	Loamy	Clay	Stony	Peat	No-Soil	

(adapted from Skalsky et al. (2008), p.12)

An additional spatial reference for the HRU data is a 30’x30’ grid (approx. 50km at the equator) spanning the globe, referred to as PX30. This coarse grid serves as an explicit spatial reference in the further organization of data and combination with climate information. The combination of PX30 spatial reference, HRU and country level administrative units results in the delineation of approximately 103,000 simulation units (SimU) (Figure 34). These SimUs are the finest spatial disaggregation within Globiom. Agricultural input data has to be provided at this level. Also yield potential data for crops provided by the EPIC crop model is aggregated on this level. However, unless such spatial detail is desirable the model is commonly solved at a coarser 200kmx200km resolution (within county boundaries) level for computational reasons and input data is averaged accordingly (Havlík et al. 2011).

Figure 34. Simulation Unit delineation



Construction of simulation units by combination of HRUs (green lines), country boundaries (black lines), and 30arcmin spatial reference (blue grid). The 5arcmin grid for input data is shown in grey (taken from Skalsky et al. (2008).

At SimU level input data for crops, livestock production and forestry is provided. For crops this data comprises current production area, area available for crop production, yield potential and input requirements such as nitrogen, phosphorous and water. Also data for forestry production and livestock production is fed into Globiom at SimU resolution by the supporting models G4M and Ruminant (Mosnier et al. 2014).

Input requirements are defined for each crop and the management systems “subsistence”, “low input rainfed”, “high input rainfed” and “irrigated”. Production functions are fixed input ratio Leontief functions but Globiom allows for shifts in the mixture of management systems in each SimU so that at SimU level changes in resource efficiency are allowed (Havlík et al. 2011).

2.4.4 Data requirements

The management system concept in Globiom is adapted from You and Wood (2006). The concept is used for the generation of the MapSpam database of downscaled current production areas. This database compiles data of crop production area from various sources at differing spatial scales. This data is then spatially disaggregated to a 5`min spatial grid using a combination of land use data, administrative data and crop suitability models (You, Wood and

Wood-Sichra 2009). Data from this database is used to provide spatially explicit crop production area as a baseline calibration reference. Such data would therefore be necessary to include coffee in Globiom. However, existing datasets of spatially explicit area data were unsatisfactory for coffee for a number of reasons (Eriyagama et al. 2014).

Land cover composition for each SimU is derived from the GLC2000 global land cover database. The GLC2000 database uses satellite data that is classified into the land cover classes and offers a globally coherent classification of land cover for the year 2000 (European Commission 2003). The different land use classes from GLC 2000 are redefined into forest land, agricultural land and herbaceous cover to define land resource endowments of SimUs.

Yield potential is modelled by the Environmental Policy Integrated Climate model (EPIC) (Williams and Singh 1995) and introduced in Globiom at SimU level. EPIC simulates major biophysical processes in agricultural systems at global scale. The model was used to assess the yield potential of 17 crops that are implemented in Globiom for three (“low input rainfed”, “high input rainfed” and “irrigated”) of the generic production system classes. To model the impacts of climate change yield potential was assessed for all SimUs with current crop area for current and future conditions. The impact data is applied as relative yield shift factors for each SimU during Globiom scenario simulations. As this data is necessary to assess the relative impacts of climate change on crop production this is necessary data to model coffee.

2.4.5 Conclusion

The feature that Globiom is spatially explicit and has previously been used to assess the impacts of climate change on food security makes this modeling framework suitable to integrate the questions raised in earlier sections of this chapter. It combines demand scenarios based on the projections of income and population that can be derived from the emission scenarios that drive climate change. Shifts in comparative advantages of supply can be modeled at a spatial resolution adequate for crop production. To include coffee in this modeling framework demand scenarios will have to be defined and spatially explicit area data and spatially explicit yield potential data are required.

2.5 Conclusion

In this chapter the necessary components of an integrated climate change impacts assessment of global coffee production were reviewed. Modeling the impacts of climate change on crop production was shown to follow a chain of models. At the top are the emission scenarios that drive the global climate models. The output of these GCMs has to be downscaled to resolutions relevant for agriculture. Biophysical models of the climate-yield or climate-crop presence relationship were reviewed here. These models are usually fitted to match current climate data, for example WorldClim data, and are then extrapolated on the future climate data. The resulting shifts in climatic suitability change the comparative advantages of regions, and thus the equilibrium in a PE framework like Globiom. This sequential modeling approach (as shown in Figure 1) will be expanded in the following chapters with the objective to conduct a globally coherent assessment of the impacts of climate change for both coffee crops.

As was demonstrated here previously published assessments of the impacts of projected climatic changes were largely regionally confined, limited to a single species (mostly Arabica) and largely stopped short of going beyond climatic suitability. Few exceptions incorporated economic indicators to estimate yield effects or affected area. No study looked at the effects of projected demand side changes.

Hence, here the objectives for this thesis were further founded in previous research. In the next chapter the first shortcoming will be addressed: a globally coherent model for both coffee species. This data will later be used to estimate yield potential for coffee production. Then the necessary data to relate these impacts to harvested area will be prepared. This allowed the inclusion of coffee in Globiom.

3 Climate change profile of global suitability for Arabica and Robusta coffee

Previous assessments of climate change impacts on coffee either used common denominators of climatic suitability for coffee production to map risk areas (Simonett 1988; Zullo et al. 2011), or used correlational approaches on temporal (Gay Garcia et al. 2006) or spatial distribution models of coffee production (Schroth et al. 2009; Davis et al. 2012). Using annual mean temperatures as a descriptor a study on Robusta in Uganda concluded that only high altitudes remain suitable (Simonett 1988). In addition to annual mean temperature Zullo et al. (2011) included water deficit and frost risk in their model to project a southward migration of Arabica production. The study by Gay Garcia et al. (2006) identified yield risks in Mexico to correlate with temperature variables and suggested that under climate change scenarios economically viable yields could be unachievable by 2020. Davis et al. (2012) concluded that areas that are climatically suitable for indigenous coffee varieties may disappear in future scenarios. Schroth et al. (2009) found a similar drastic impact on Mexican coffee production associated with increasing seasonal temperatures.

Thus, previous studies on climate change impacts demonstrated possible drastic impacts on coffee cultivation: Latitudinal migration, altitudinal migration or complete abandonment of coffee production. However, results were limited to local levels and global trends remained unclear.

As discussed in chapter 2.2 several studies investigated climate change impacts *on Coffea arabica* using the Maxent species distribution modeling software (Phillips et al. 2006) that is

Authorship:

The idea to apply the Maxent modeling software on global scale for *C. arabica* was first conceived by Peter Läderach (International Center for Tropical Agriculture (CIAT)) and published in Ovalle-Rivera et al. (2015). The extension of this approach by using various parameter combinations and alternative algorithms in a model ensemble was my own work (Bunn et al. 2015).

Part of the occurrence data used for model training was prepared by Oriana Ovalle-Rivera (International Center for Tropical Agriculture (CIAT) as presented in Ovalle-Rivera et al. (2015).

based on machine learning concepts. It has been suggested that such models are prone to overfitting to a biased representation of suitable climate when not used with appropriate parameter choices, the resulting reduction in suitable area in the present studies could thus be model inherent.

Defining appropriate parameter values for robust models is the topic of a vast body of literature. In lieu of reliable data that allows the comparison of intertemporal climatic and species distribution changes there is no clear guidance for parameter values that allow reliable extrapolation (Elith and Graham 2009). This is despite uncertainties arising from the distribution modeling that are larger than those from GCM outputs (Diniz-Filho et al. 2009).

To overcome this limitation it was suggested that ensemble outputs of various models are more robust and allow for explicit uncertainty analysis (Araujo and New 2007, Diniz-Filho et al. 2009). For probabilistic model outputs from overconfident models it was demonstrated that multi-model ensemble means improve prediction skill (Weigel, Liniger and Appenzeller 2008). Such ensemble approaches have been employed to investigate the potential indirect land use change effects on ecosystems by the migration of viticulture (Hannah et al. 2013) and to generate risk maps of Dengue fever (Bhatt et al. 2013).

The objective of this chapter is to predict current and future climate suitability for coffee (Arabica and Robusta) production on a global scale. The distribution of suitability under current and future conditions is compared and a global impact profile of climate change on coffee production is derived. An ensemble approach is chosen to improve the robustness of the analysis over previous studies.

First the assembly of a comprehensive global dataset of known occurrence locations of either coffee species is described. On this data three popular machine-learning algorithms (Support Vector Machines (Karatzoglou, Meyer and Hornik 2006), Random Forest (Breiman 2001) and MaxEnt (Phillips et al. 2006) using distinct parameter combinations as outlined below were trained, resulting in a total of 135 models. The model performance was evaluated against the performance of a trivial inverse distance based model. Finally, the models were extrapolated on interpolated climate data of current and future climatic conditions and the mean suitability score for each global pixel cell was derived. The future climate data was generated by downscaling GCM models for the RCP 2.6, RCP 6.0 and RCP 8.5 emission pathways. Impacts were analyzed by latitude, altitude, regions and land-use classes to hypothesize future impact scenarios on global coffee production.

3.1 Methodology – Machine learning model ensemble

The guiding principle to develop the methodology was to choose model parameters such that the resulting model generalizes well. In previous studies parameter choices and the use of a single classification algorithm likely caused some of the reported negative impacts of climate change. Here, input data, algorithm parameterization and the ensemble approach itself were designed to avoid such overfitting.

Three types of data were employed for the classification of climate as suitable or unsuitable: (i) Locations of known current occurrences specify suitable climates. (ii) Random background samples from the environment were needed for a comparison of the climate at the ocurrence locations and other climatic conditions in the study region and, of course, (iii) spatial climate variables that are relevant for bioclimatic suitability. First the assembly of a comprehensive database of occurrence locations of *Coffea arabica* and *Coffea canephora* is described, then the generation of random background samples will be discussed.

3.1.1 Current occurrence data

Occurrence location data locate climates that are currently suitable for coffee production. The occurrence points in the database were derived from three principal sources: (i) Geo-referenced coffee farms, (ii) geo-referenced coffee producing municipalities in Brazil and (iii) geo-referenced coffee growing areas identified from satellite data (google earth) where data source (i) and (ii) were not available.

The majority of the occurrence locations originate from a database of more than 62.000 geo-referenced individual coffee farms with predominantly *Coffea arabica* and some *Coffea canephora* from all over the world. This database was developed during several regional projects that were conducted by the International Center for Tropical Agriculture (CIAT) in collaboration with coffee cooperatives and cooperating research organizations.

Unlike the *C. canephora* data, the majority of *C. arabica* locations was not recorded for modeling purposes. This data was thus highly clustered in the project regions. To avoid a biased representation of the respective climatic conditions in these regions the database had to be stratified. The initial database for *C. arabica* was therefore subsampled by conducting a principal component analysis using the 19 bioclimatic variables from WorldClim to identify typical climates. From each climate cluster a random representative sample was picked. This reduced the original sample to 1772 unique presence locations for *C. arabica* (Ovalle-Rivera et al. 2015).

The resulting database did not include all of the dominant growing regions in Brazil, where 36% of global Arabica production is based (USDA 2012). Therefore data provided by IBGE (2012) was used to identify municipalities that are characterized by at least 75% of production being one of the two species. These municipalities were geo-referenced as suitable locations for either one of the respective species as follows:

Production statistics were available on municipal level but are aggregated over Robusta and Arabica coffee (IBGE 2012). Thus, this information was not directly usable as the interest here was the differential impact of climate change on the two species. For 2006, however, census data for large commercial farms differentiated between Robusta and Arabica production. The two datasets were not consistent for all regions. Therefore the time series data was split according to shares of production calculated from the census data.

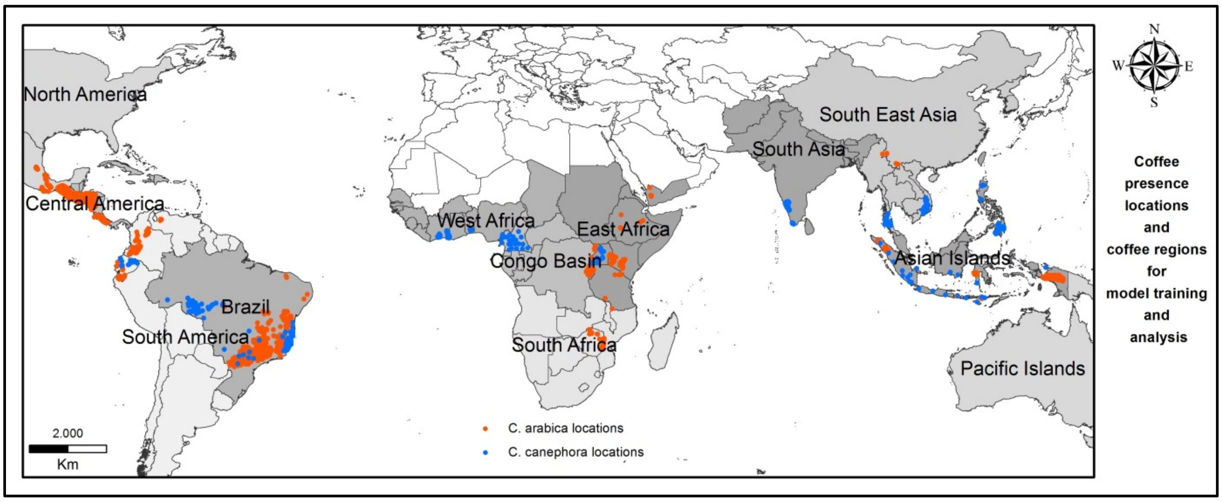
First, the total production for the `95-`04 period was averaged as this dataset also included small producers. Municipalities were then separated into 5 percentiles, with 4365 of 5490 municipalities not producing coffee. In ArcGis 10.1 random reference points were generated as specified by the percentile membership of the municipality: No point for municipalities without significant production and five points in municipalities of the highest production percentile and so on. Occurrence locations were confined to agricultural area as defined in the GLC2000 dataset (European Commission 2003).

These locations were then divided into Robusta occurrence points and Arabica occurrence points based on the production share that was calculated from the census data, e.g. if according to the census data 80% of production was Arabica and 20% Robusta, the points were assigned accordingly, rounding down. To reduce spatial autocorrelation a minimum distance between points of 0.05 decimal degrees was specified. The resulting set of occurrences consisted of about 1300 individual presences of which about 350 belonged to Robusta and 950 to Arabica.

A comprehensive set of presence records in all coffee producing regions is desirable so that all suitable climates are represented in the database (Elith et al. 2011). Therefore the database that resulted from the combination of data from CIAT projects and Brazilian municipalities was supplemented by generating additional occurrence points using publicly available information about the distribution of coffee production. Where available, satellite imagery was used to identify precise locations in regions that are known for coffee production and were not represented in the database.

All occurrences were reduced to unique occurrence pixels at 5arcmin resolution. The resulting occurrences dataset included 2861 unique pixel cells for *Coffea arabica* distributed over 26 countries that together accounted for 95% of global Arabica output in the period '98-'02 (USDA 2012). For *C. canephora* the presence dataset included 364 unique pixel cells distributed over 11 countries that together accounted for 92% of global Robusta output in the period '98-'02 (USDA 2012) (Table 8). Figure 35 shows the distribution of the coffee occurrence locations in the database. Also shown are major geographic regions of coffee production that were used to analyze impacts.

Figure 35. Worldwide coffee occurrence locations and coffee regions



Blue dots indicate *C. canephora*, red dots *C. arabica* locations; Grey shading shows coffee region definitions used to analyze the regional distribution of impacts (own data and representation).

Table 8. Distribution of occurrence locations for model training

Country	Arabica		Robusta	
	Locations No.	Production share in %	Locations No.	Production share in %
Benin	0	0	5***	0
Brazil	1042**	36	164**	16
Burundi	8***	1	0	0
Cameroon	0	0	22***	2
China	5***	0	0	0
Colombia	169*	16	0	0
Congo (Dem. Rep.)	1***	0	0	0
Costa Rica	124*	4	0	0
Côte d'Ivoire	0	0	20***	10
Ecuador	43*	1	17*	1
El Salvador	55*	3	0	0
Ethiopia	9***	5	0	0
Guatemala	230*	6	0	0
Honduras	158*	4	0	0
India	2***	3	21***	7
Indonesia	184*	1	22***	16
Kenya	111*	2	0	0
Mexico	407*	7	0	0
Mozambique	14*	0	0	0
Nicaragua	170*	2	0	0
Panama	4***	0	0	0
Philippines	0	0	20***	2
Rwanda	18*	0	0	0
Tanzania	9***	1	0	1
Thailand	0	0	15***	3
Uganda	24*	1	17*	7
United States	2***	0	0	0
Venezuela	3***	2	0	0
Vietnam	19*	0	41*	27
Yemen	4***	0	0	0
Zimbabwe	46*	0	0	0
Sum	2861	95	364	92

Locations (*) from projects, (**) survey data (IBGE 2012), and (***) geo-referenciation and shares of global production '98-'02 (USDA 2012).

3.1.2 Variable choice

While on local scale soil quality, aspect, or local climate dynamics are important, on global scale broad climate variables are decisive for suitability. Therefore, such variables were used for site classification. For variable choice there are two schools of thought, one that includes all available variables in the model and lets algorithms decide which variables to pick, and one that only includes variables that are deemed significant by the modeler (Dormann et al. 2013).

Often true biophysical dependencies are unknown so that the inclusion of several variables offers the advantage that the resulting model is not only based on statistical significance but is also less specific, as all variation within the dataset is included. However, in cases where one or more variables are collinear (i.e. a variable may be linearly constructed from another variable, thus one variable will partly explain the effect of the other) statistical models are unable to differentiate the underlying effects. Extrapolation of such models to geographical spaces or climate conditions with a different collinearity pattern may thus yield erroneous results (Thuiller 2004).

The inclusion of a limited set of climate variables based on pre-existing knowledge about biophysical effects avoids the latter pitfall, but nevertheless has its own problems. All real world data is collinear to some extent because they derive from underlying processes that may be unmeasurable. In such cases statistical methods may be unable to estimate true effect sizes. Furthermore, by removing variable information from the dataset the resulting model may not reflect the entire biological range, thus extrapolation may overestimate effects of changes in variable patterns (Harris et al. 2013). For machine learning based classification of current and future climatic suitability overfitting has been identified as a problem. Therefore, for the mapping of current and future coffee production areas the entire environmental range should be included and all available variable data was employed.

Data for the current climate (1950–2000) was downloaded from the WorldClim global climate data base on 2.5 arcmin resolution (Hijmans et al. 2005). The dataset provides interpolated climate layers for 19 bioclimatic variables based on historical data (Table 1). These variables represent patterns found in monthly weather station data, e.g. annual temperature and precipitation extremes, seasonality and means. The interpolated climate surfaces for global land areas were generated from a comprehensive set of climate data sources for the globe. Especially regions with low population density are underrepresented in the station data (comp. chapter 2.1.1.1).

Future climate data was projected by 5 GCMs (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, MIROC-ESM-CHEM, and NorESM1-M[6]) from the IPCC 5th assessment report. These GCMs were chosen because they are representative of projected global mean temperature and precipitation changes across AR5 models (Warszawski et al. 2014). The native outputs of the GCMs were downscaled using the delta method as in (Ramirez and Jarvis 2010): The

[6] Descriptions of the individual GCMs can be obtained from Stocker et al. (2013)

difference between model outputs for current conditions and the average for the period 2040 to 2069 was computed. The resulting layers were smoothed to a 2.5arcmin resolution and applied to the WorldClim layers for current climate. The result was a bias corrected high resolution surface for the 2050s of the same 19 bioclimatic variables as for current climate. Three representative concentration pathways with low (RCP 2.6), intermediate (RCP 6.0) and high emissions (RCP 8.5) were considered.

3.1.3 Algorithm choice

Previously, the most widely used algorithm to assess the impacts of climate change on coffee was Maxent. This algorithm of the machine learning class was complemented here with two additional classification methods to improve the performance of the resulting model ensemble. Thus, three popular machine learning algorithms were used: Maxent, Support Vector Machines (SVM) and Random Forest. Maxent (Phillips et al. 2006) is one of the most popular species distribution modeling software in ecology (Merow et al. 2013). SVM is one of the most popular general purpose classification algorithms. The implementation in the R package "kernlab" (Karatzoglou et al. 2006) was used here. Random Forests (Breiman 2001) are increasingly popular and have been shown to be useful in ecology (Prasad, Iverson and Liaw 2006). The randomForest package (Liaw and Wiener 2002) in R was used. For all computations the open source software R was used (R Core Team 2014).

3.1.4 Background locations

In order to fit a function that describes suitable climates the classification algorithms compare variable patterns found at occurrence locations with the pattern found in potentially suitable environments (Figure 17). To characterize these environments random background samples were taken from locations that are not known locations of present occurrence.

The background samples were chosen such that both trivial classification and overtraining of the algorithms were avoided. In ecology, a trade-off persists between predictive performance and generalization capability. For example, a model that always correctly separates known presence locations from the random background samples may be an undesirable model. This is because it underestimates the true environmental range in cases where the known presences incompletely represent the true distribution. However, a more general model that would also correctly predict unknown presence locations may overestimate the environmental range.

No optimization framework for the definition of background parameters and modeling approaches exists to date (Elith and Graham 2009). Therefore rather than using a single sampling strategy a model ensemble is used here. Several background sampling parameters were used that were within reasonable ranges for the geographical extent from which the background sample is drawn (3.1.4.1), and the number of samples (3.1.4.2). Furthermore, the remaining sampling bias in the occurrence location database was accounted for using the biased background sampling method (Dudík, Phillips and Schapire 2005).

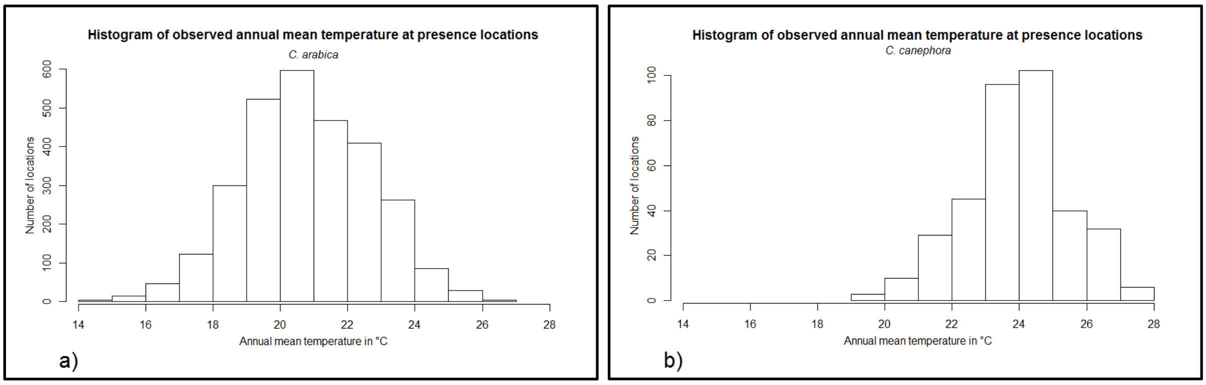
3.1.4.1 Background extent

The first important factor in background sampling is the extent of the area from which locations may be picked. If the extent is too limited the algorithms are unable to distinguish patterns that act on relevant scales, but if too broad the recognized patterns may be trivial. The choice of the geographic extent should therefore reflect prior knowledge of the species distribution and be adequate to the geographical scale of the study (VanDerWal et al. 2009).

In the literature three basic forms of choosing the extent of the study area can be found. Frequently, the motivation for a study includes a political component such as the distribution of a species within a country or a wildlife reserve, the study extent is thus a political one. Also common is the geographic notion that things that are closer are also more similar, defining the study area based on distance to presence points. And last, some authors demand that the study area should be defined by some biological prior knowledge about the studied species.

To reflect this disagreement three different background concepts were employed: a political one, a biophysical one and a geographic one. The first background was defined over all Robusta or Arabica producing countries (USDA 2012, ICO 2013), the second by limiting the environment to the observed spread of annual mean temperature for each species location sample (*C. arabica*: 14°C – 26.4°C; *C. canephora* 19.2°C – 27.8°C annual mean temp.; Figure 36); and the latter by using a 4.5° buffer around presence locations (approx. 500km at the equator in degrees).

Figure 36. Histogram of annual mean temperature at coffee occurrence locations



a) *C. arabica* and b) *C. canephora* (own data and representation).

3.1.4.2 Background ratio

The number of background samples has a great effect on model accuracy. There is a balance between too many and too few background samples, leading to under- or over estimation of absence areas. The literature agrees that the ratio should be at least 1:1 (Barbet-Massin et al. 2012). Too few background samples do not allow for a clear distinction between presence and background, commonly leading to an over-prediction of distribution, while too many result in under-prediction. In extreme cases, e.g. a ratio of 20:1, the background class may always be predicted while accepting the under-prediction of presence locations as error.

Therefore sampling ratios 1:1, 2:1, 4:1, 6:1, 8:1 of background to presence points were chosen.

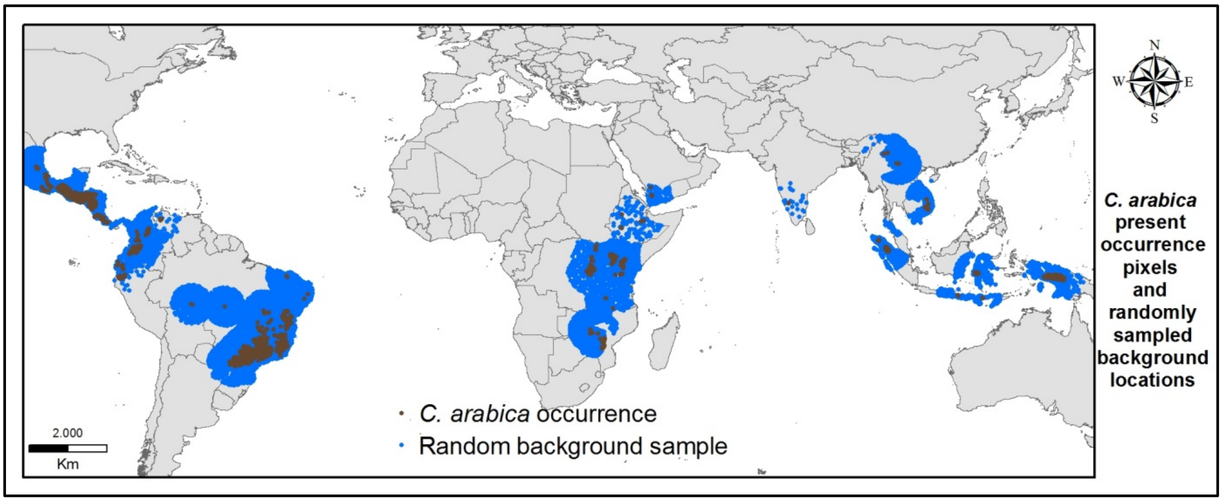
3.1.4.3 Biased background method

To account for remaining distributional bias in the presence data the biased background sampling method was used. Using this approach the geographic bias in presence locations is cancelled out by a similar bias in the random background noise samples (Dudík et al. 2005). To assess the bias in the presence sample a measure was calculated of how well the presence location database is representative of actual coffee areas: the share of total presence points present in a country divided by the share of harvested area of global area in the country.

This way, for each country a prior probability that its pixel cells are picked during background sampling was assigned. Thus, pixel cells that are in countries that were overrepresented were more likely to be picked as a background cell. This was meant to avoid the overtraining of the models with patterns from regions that were overrepresented in the occurrence location database.

The resulting background sample is shown for one example in Figure 37. In this example, the background samples have an 8:1 ratio and were taken from an extent that was limited to a distance of 4.5° to presence samples. The varying density of samples reflects the biased background method.

Figure 37. Distribution of presence and background sample of *C. arabica*.



Brown dots indicate presence locations; blue dots indicate background locations using an 8:1 ratio, a geographically defined extent and biased background sampling (own data and representation).

3.1.5 Algorithm parameterization

Most machine learning algorithms include a regularization parameter that allows the user to adjust a trade-off between optimal model fit and generalization. Optimal parameter values are usually dependent on the characteristics of the provided data. Therefore initially relevant parameter spaces were defined by conducting a grid search over feasible parameter levels.

Maxent and SVM explicitly penalize variables that do not add information to the fitted model and only add model noise. The random forest algorithm is inherently designed to avoid overfitting by constructing an ensemble of several individual decision trees that at each decision node pick variables at random. It is therefore claimed that overfitting does not occur due to the independence of the individual trees. Nevertheless, the chance that meaningless variables influence the model can be reduced by increasing the number of randomly picked variables at the nodes (Breiman 2001).

3.1.5.1 Regularization parameters

In support vector machines the cost parameter C penalizes misclassification of data. Misclassified points are assigned a cost to the distance with increasing distance to the classification hyperplane. The default setting is “1”. Decreasing values decrease the cost of misclassification.

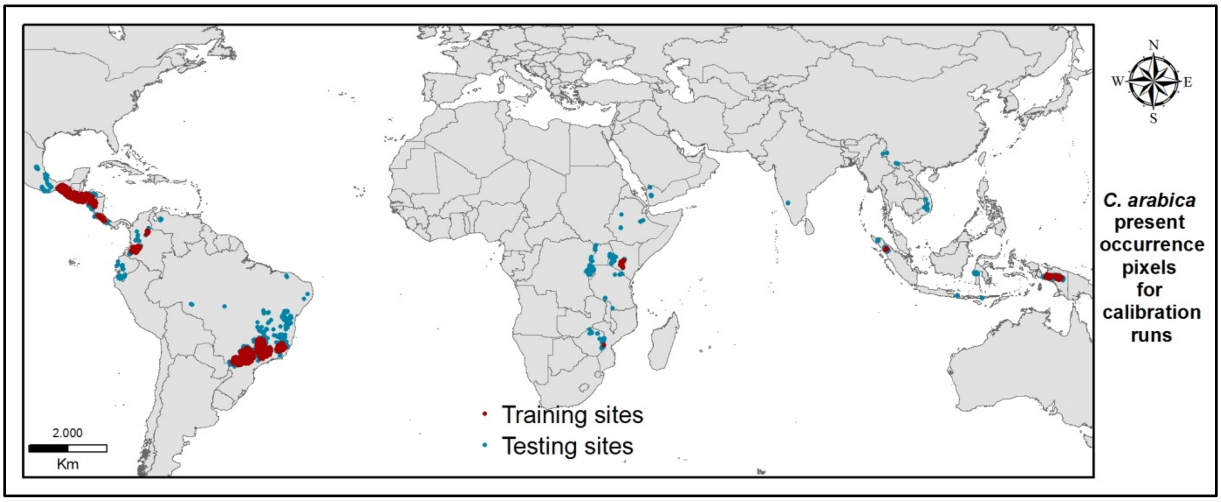
The regularization parameter β in Maxent trades off model complexity and variable importance. The regularization parameter both ensures that model constraints are not enforced too strictly to avoid overfitting. It also has an effect on variable selection by penalizing the use of additional variables. Choosing a higher regularization parameter therefore results in a model with less variables, and the included variables are less enforced. The default value is “1” which is generally seen a reasonable value for multi-species studies, though at the lower end of the range of recommended values.

Random Forests were introduced by (Breiman 2001). A major advantage is the robustness against overfitting with increasing tree number. The R implementation by (Liaw and Wiener 2002) is used here. Because the classification is averaged over a large number of trees, increasing tree number decreases the risk of overfitting. It is therefore recommended to grow as many trees as computationally feasible. Another key parameter is the number of variables picked at random at each tree node to be evaluated. The default is the square root of the variable number. However, if there are many variables that do not add information a too small number increases the risk of adding unwarranted model complexity. On the other hand, picking few variables may result in more accurate models as also variables that account for little variation are included (Boulesteix et al. 2012). We vary the number of variables picked at each node from 1 to 19, and compare using 200 and 1500 trees. However, in the calibration run increasing tree number did not improve generalization above 1000 trees; picking less variables at each node had a larger effect on generalization. Thus, the number of variables picked at each node was varied from 2, 4, and 8 using 1000 trees.

3.1.5.2 Calibration test data preparation

To define regularization parameter values a grid search was conducted for each algorithm on reasonable values derived from the literature that were meant to generalize the model. The generalization efficacy was determined by training on two thirds of the points and testing on the third of the points that is geographically distant from the training data, comparing AUC values on this dataset. The occurrence point database was divided into a training dataset and a testing dataset as follows: for each point the distances to all other points were calculated and the lowest percentile defined (with 2861 points that would be the distance to the 28^{th} nearest point). From this the 25% of points were picked as a test data set that were furthest away from other points. The result was a data set of test points that is spatially distant from the training points (Figure 38).

Figure 38. Distribution of *C. arabica* locations used for model calibration

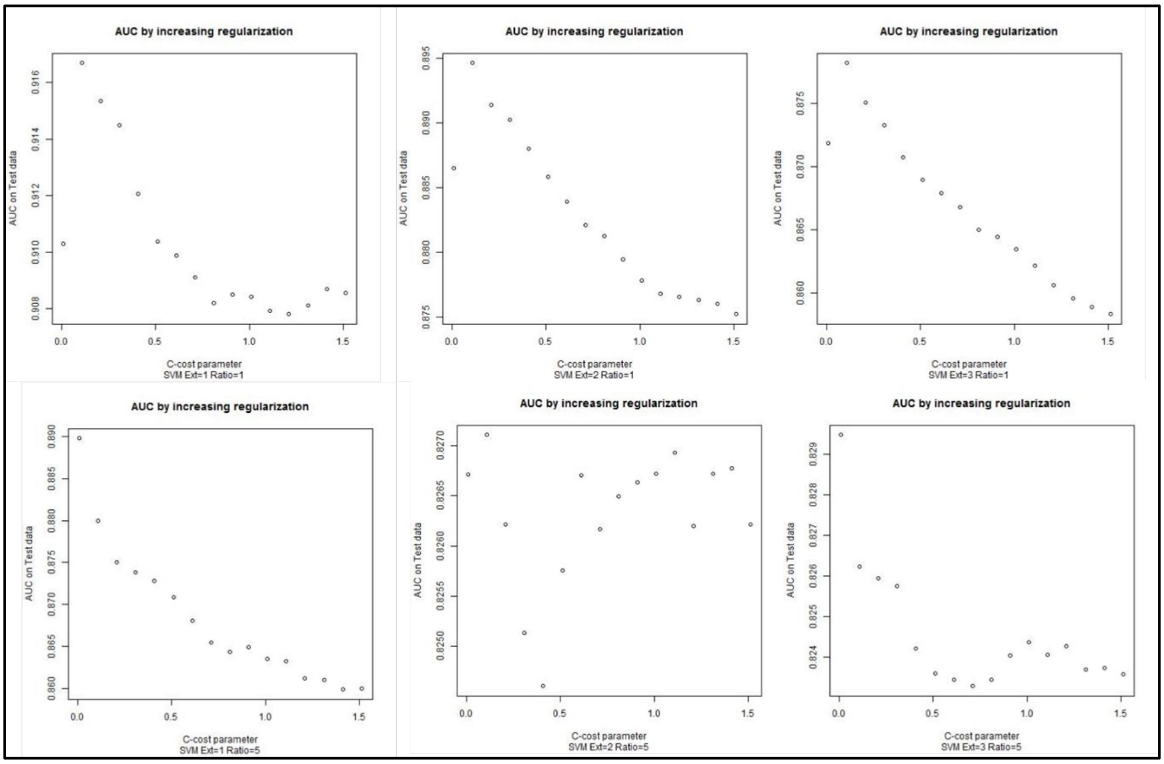


Training sites in red and testing sites in blue (own data and representation).

3.1.5.3 Final parameter choices

Exemplary Figure 39 shows the AUC results for the grid search over the C-cost parameter for six extent-background sample number combinations. Generally, a decreasing C-cost parameter increased the AUC on the test data set, indicating a better generalization.

Figure 39. Decreasing prediction of test sites with increasing SVM C-cost parameter



AUC result on test sites of grid search for various extent-background sample combinations for SVM C-cost parameter (own data and representation).

Based on these grid searches three values per algorithm were picked that improve model generalization compared to default settings. The final values are presented in Table 9.

Table 9. Classification algorithm parameter choices after grid search

Algorithm	Maxent	SVM	Random Forest
Parameter	β regularization	C cost parameter	Variables at node
Very general	20	0.05	2
Intermediate	5	0.5	4
Fitted	0.01	1	8

(Own data)

3.1.6 Model evaluation by AUC and Kappa

To assess the performance of the individual models four measures were used: The threshold independent area under the receiver characteristic curve (AUC), a calibrated AUC measure, the threshold dependent Kappa and a calibrated Kappa.

The AUC is the standard method of model evaluation method in predictive distribution modeling as it does not require a subjective decision on defining a threshold between presence and absence of the continuous probability model output data. It summarizes the ranking of presence points versus the ranking of background samples. If all presence sites have a higher value than background sites its values is 1, while a value of 0.5 should reflect a model that is no better than chance. However, it has been criticized to be misleading when different background samples are drawn from different background extents: low predictions on geographically distant locations are trivial and may thus inflate the statistic (Lobo, Jiménez-Valverde and Real 2008).

Taking into account spatial autocorrelation of climate patterns such an effect is to be expected. (Hijmans 2012) proposed to account for such effects by means of a simple null model based on the distance to training presence locations, and calibrate the model AUC with the null model AUC. Therefore the standard AUC is supplemented by a calibrated cAUC as proposed by (Hijmans 2012) (Eq. 2):

$$cAUC = AUC + 0.5 - \max(0.5;\, nAUC) \qquad \text{(Eq.2)}$$

where nAUC is the null model AUC.

Another issue with the AUC measure is its equal weighting of omission and commission error. An increasing background to presence point ratio could thus result in a reduced AUC despite equal predictive performance. Cohen's Kappa has been proposed to be a suitable classification model evaluation statistics because it makes full use of information on errors (Fielding and Bell 1997) to assess whether a model is better than chance. Lacking true absences the kappa was calculated based on a 2.5% omission error threshold to evaluate how well the model differentiated occurrence and background.

Again, the kappa was supplemented with a calibrated kappa that was calculated compared to the null model kappa (Eq. 3):

$$cKappa = Kappa - nKappa \qquad \text{(Eq. 3)}$$

where nKappa is the null model Kappa.

To evaluate whether all parameter choices introduce significant variability an Anova analysis was conducted on the evaluation metrics with the model setting parameters as independent variables (Eq. 4):

$$\begin{gathered} Metric \sim algorithm\ choice\ (df = 2), \\ background\ extent\ (df = 2), \\ background\ to\ presence\ point\ ratio\ (df = 4), \\ regularization\ level\ (df = 2);\ N = 1350 \end{gathered} \qquad \text{(Eq. 4)}$$

Variable importance was estimated by computing AUC on each predictor variable individually using the Caret package in R (Kuhn 2008). This method applies cutoffs to the predictor data and then calculates sensitivity and specificity for each cutoff to calculate the AUC. The AUC is then a measure for variable importance.

3.1.7 Assessing the impacts

The three algorithms were trained using the above described parameter spaces: For each of the three extents five different background to occurrence ratios were employed and three regularization choices. Thus, a total of 3*3*5*3= 135 distinct models per species were trained. The trained and tested models were extrapolated on raster data for the 19 bioclimatic variables from WorldClim and for the 2050s period. This produced maps of continuous scores whether a pixel cell belongs to the absence or presence class. This is equivalent to rating each global pixel cell's climate as suitable or unsuitable for coffee production. Individual model outputs were normalized to scores from 0 to 1 and averaged for each baseline and emission scenario. To define a threshold between probabilities that represent marginal suitability and relevant suitability values we chose the lowest value at a presence location. Only pixel cells that had suitability values above this threshold were included in the analysis. Impacts were compared across latitude and altitude classes. The suitability score was summed up across 1° latitude classes and 100m altitude classes. Regional analysis of impacts was done for 12 regions of coffee production (Figure 35).

The GLC2000 global land cover database (European Commission 2003) was used to partition suitability changes to land with forest cover (GLC2000 global categories 1-9, Table 10) and land without forest cover and agricultural land (GLC200 global categories 10-18, Table 10).

Table 10. Global land use class definitions from GLC 2000

Class #	GLC Global Class (according to LCCS terminology)
1	Tree Cover, broadleaved, evergreen LCCS
2	Tree Cover, broadleaved, deciduous, closed
3	Tree Cover, broadleaved, deciduous, open
4	Tree Cover, needle-leaved, evergreen
5	Tree Cover, needle-leaved, deciduous
6	Tree Cover, mixed leaf type
7	Tree Cover, regularly flooded, fresh water (& brackish)
8	Tree Cover, regularly flooded, saline water,
9	Mosaic: Tree cover / Other natural vegetation
10	Tree Cover, burnt
11	Shrub Cover, closed-open, evergreen
12	Shrub Cover, closed-open, deciduous
13	Herbaceous Cover, closed-open
14	Sparse Herbaceous or sparse Shrub Cover
15	Regularly flooded Shrub and/or Herbaceous Cover
16	Cultivated and managed areas
17	Mosaic: Cropland / Tree Cover / Other natural vegetation
18	Mosaic: Cropland / Shrub or Grass Cover
19	Bare Areas
20	Water Bodies (natural & artificial)
21	Snow and Ice (natural & artificial)
22	Artificial surfaces and associated areas
23	NoData

Categories marked in "green" were treated as "Forest covered", and those marked in blue as "Open land" during analysis, grey categories were excluded (own representation of data from European Commission (2003).

Tropical forests are providers of diverse ecosystem services, are more species rich and hold higher carbon stocks than coffee plantations (De Beenhouwer, Aerts and Honnay 2013). However, coffee plantations are often more biologically diverse than other agricultural land (Moguel and Toledo 1999) and hold relatively high carbon stocks (van Rikxoort et al. 2014). Therefore, a conversion from natural forest to coffee plantations would result in a negative environmental impact. On the other hand, this would not necessarily be the case in a conversion from open land to coffee plantations.

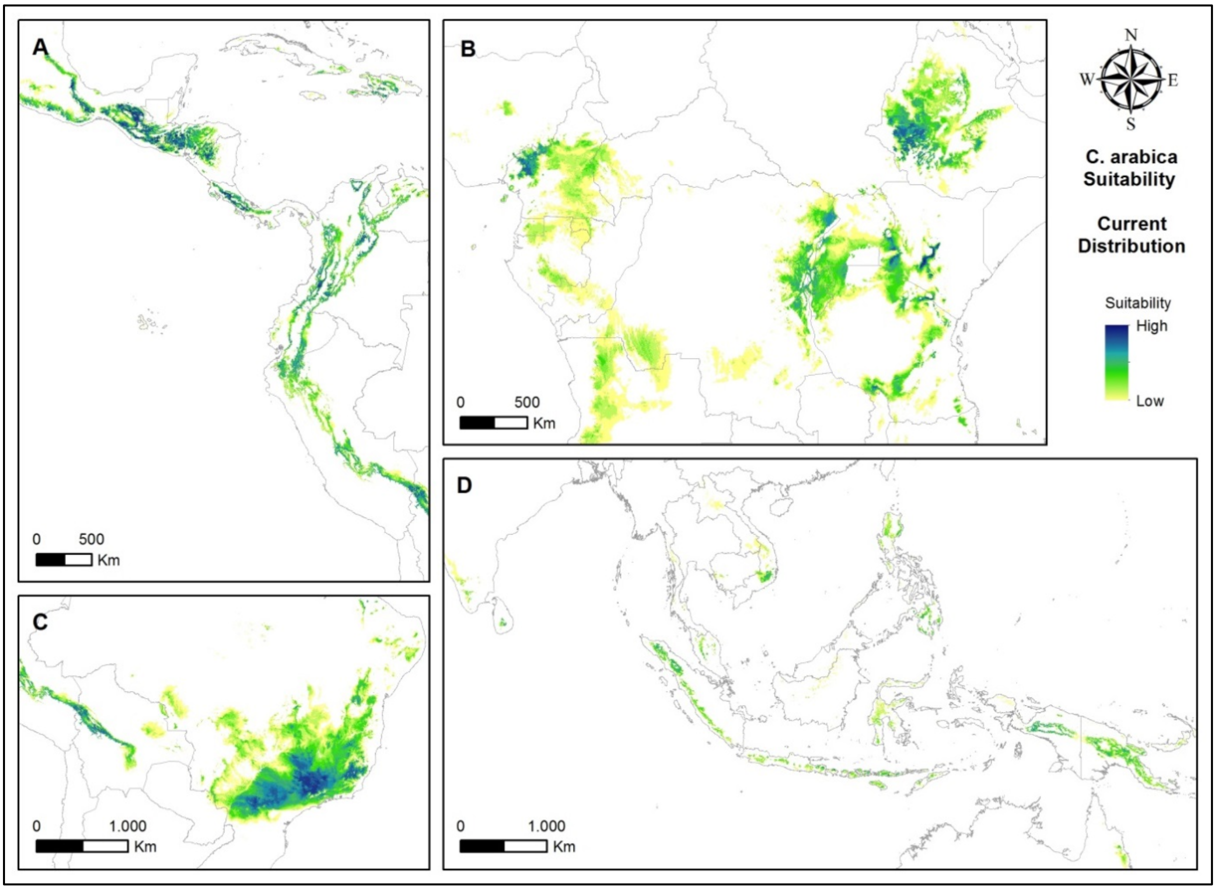
3.2 Results – The impacts of climate change on the distribution of coffee suitability

: First the spatial distribution of coffee production under current and future conditions is presented and compared. Then the values of AUC and Kappa will be shown to demonstrate the model validity. Analyzing the weight given to the variables supported the understanding of how climate change will change the distribution of coffee suitability.

3.2.1 Current coffee suitability

A global map of current suitability for coffee production resulted from the extrapolation of the classification models on raster data for the bioclimatic variables from WorldClim (Figure 40).

Figure 40. *Coffea arabica* current suitability distribution in major production regions

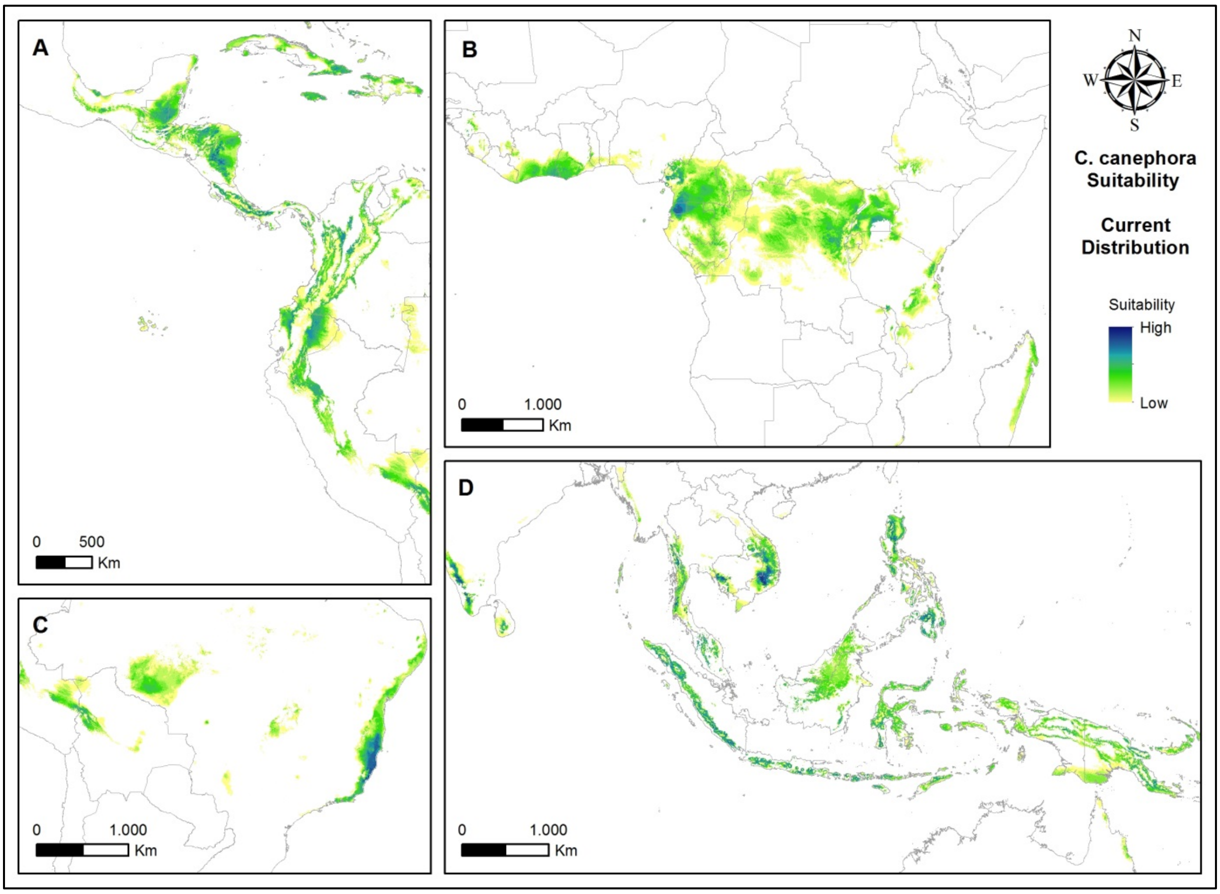


Shown is the mean suitability score of the model ensemble; dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (own data and representation).

The largest areas suitable for *C. arabica* based production can be found in the Brazilian Minas Gerais province. Other highly suitable areas are located in Central America and the Ethiopian highland region. Madagascar was also found to be highly suitable despite not being a major producer today. Other African origins, and Asian origins, were rated as predominantly of intermediate climatic suitability for Arabica production (Figure 40).

Larger areas highly suitable for *C. canephora* are in the Brazilian Espirito Santo region, West Africa, the lower regions of Central America and in mountainous locations in Asia, especially the Philippines, Indonesia and Vietnam (Figure 41).

Figure 41. *Coffea canephora* current suitability distribution in major production regions

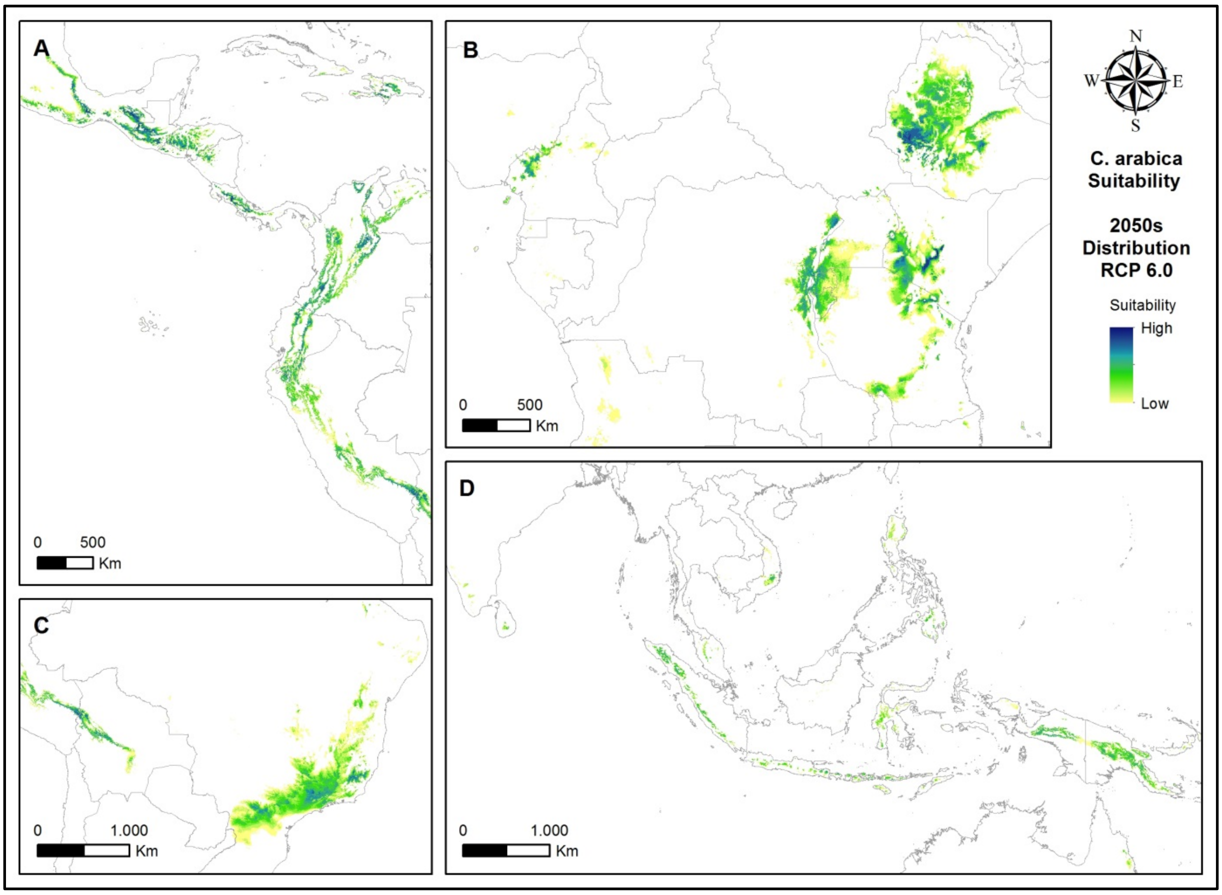


Shown is the mean suitability score of the model ensemble; dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central and West Africa, C) Brazil, D) Asia (own data and representation).

3.2.2 Future coffee suitability

Application of the same 135 models to downscaled GCM outputs resulted in maps of suitability scores for the 2050s period. Figure 42 shows the distribution of suitability for Arabica coffee in the RCP 6.0 scenario. In comparison with the current distribution area suitable for Arabica production in Brazil and Central America is reduced. In Asia and Africa the reduction appears to be less drastic. Additional maps for the RCP 2.6 and 8.5 scenarios can be found in Annex 2 (Future coffee suitability).

Figure 42. *Coffea arabica* 2050s suitability distribution in the RCP 6.0 scenario

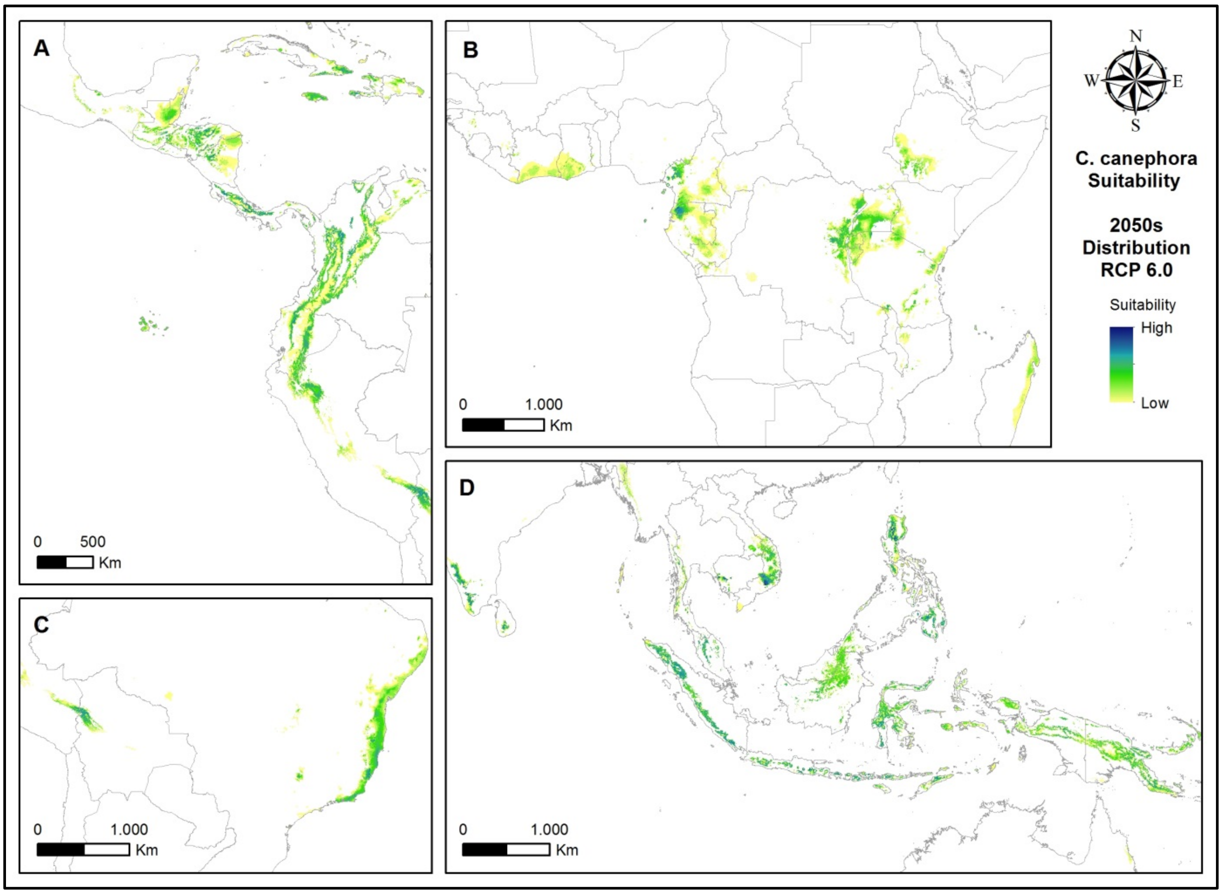


Shown is the mean suitability score of the model ensemble; dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (own data and representation).

Figure 43 shows the resulting distribution of suitability for Robusta coffee production in the same period and scenario. Again, additional maps for RCP 2.6 and RCP 8.5 can be found in Annex 2 (Future coffee suitability).

The largest difference in area to current conditions can be observed in the Congo basin where currently extensive areas could be used for Robusta production under baseline climate conditions but not by the 2050s. Suitability for Robusta in Asia appears to migrate to different locations.

Figure 43. ***Coffea canephora* 2050s suitability distribution in the RCP 6.0 scenario**

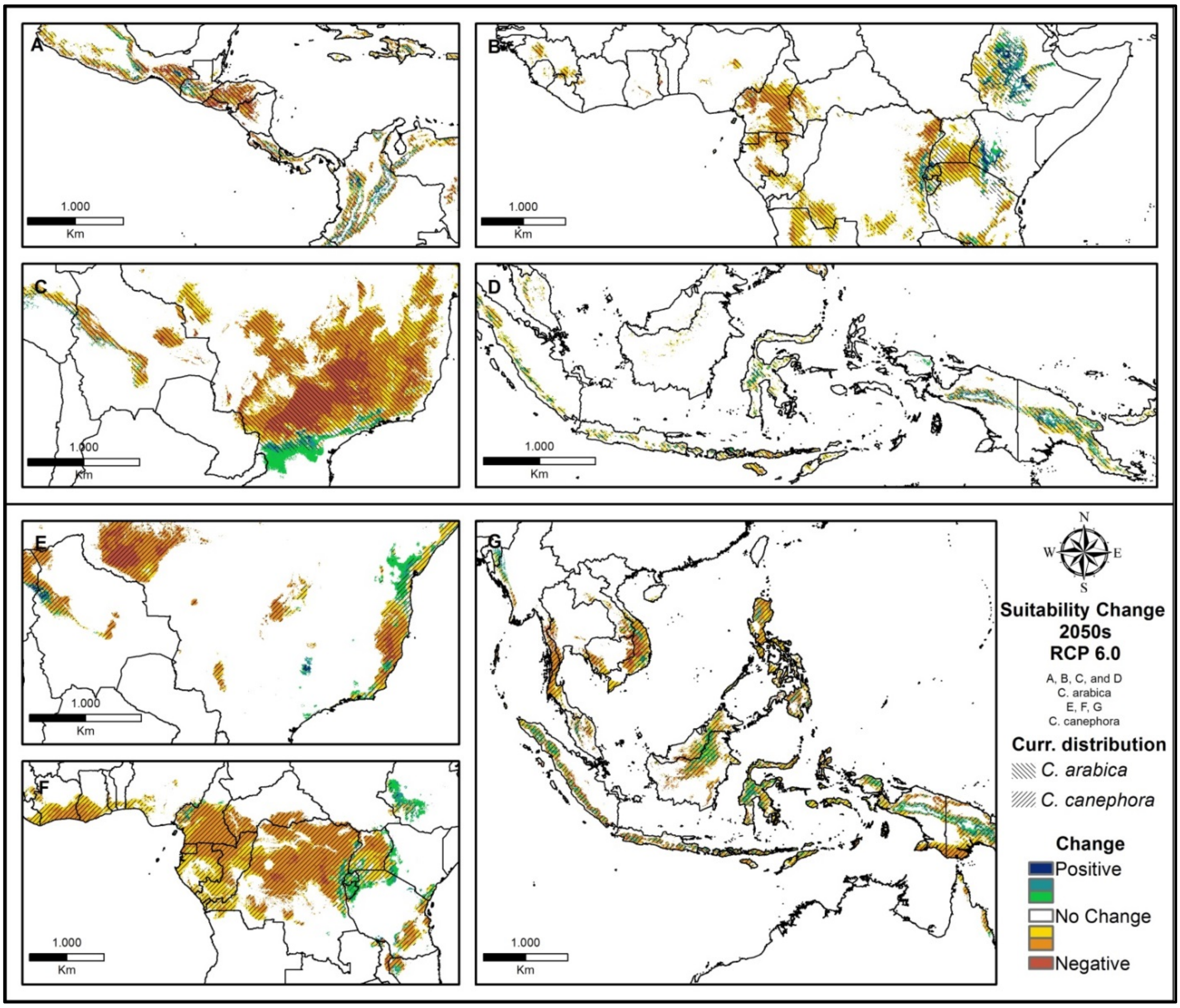


Shown is the mean suitability score of the model ensemble; dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (own data and representation).

To assess the changes in future climatic suitability the difference between current and future (2050s) mean suitability scores were calculated. For the RCP 6.0 scenario the suitability change is shown in Figure 44 (A-D) for the dominant production regions of *C. arabica* in Brazil, Latin America, Asia and the species origin in East Africa. The results indicated losses of suitability in the Brazilian production regions with possible positive changes at its Southern margin. In the rest of Latin America positive suitability changes could be observed in higher altitudes than previous production. In East Africa the mean of models indicated positive suitability changes especially in Ethiopian, Ugandan and Kenyan high regions. In Indonesia and the Philippines a pattern of altitudinal migration similar to South American locations was projected.

Figure 44 (E-G) shows the changes in suitability for *C. canephora* production by 2050 in the RCP 6.0 scenario GCM outputs for Brazil, its region of origin West Africa, and the most important region of Robusta production in South East Asia and the Asian island states. The Brazilian states of Rondonia and Espirito Santo will see stark losses of suitability. Decreasing suitability dominates in the Congo basin and coastal regions of West Africa. Higher altitudes along the equator are most likely to see suitability increases. In South East Asia the dominant Vietnamese production regions were shown to lose suitability, while in Indonesia and the Philippines higher altitudes will become more suitable for Robusta production. Maps for the RCP 2.6 and RCP 8.5 emission scenarios can be found in Annex 2 (Future coffee suitability).

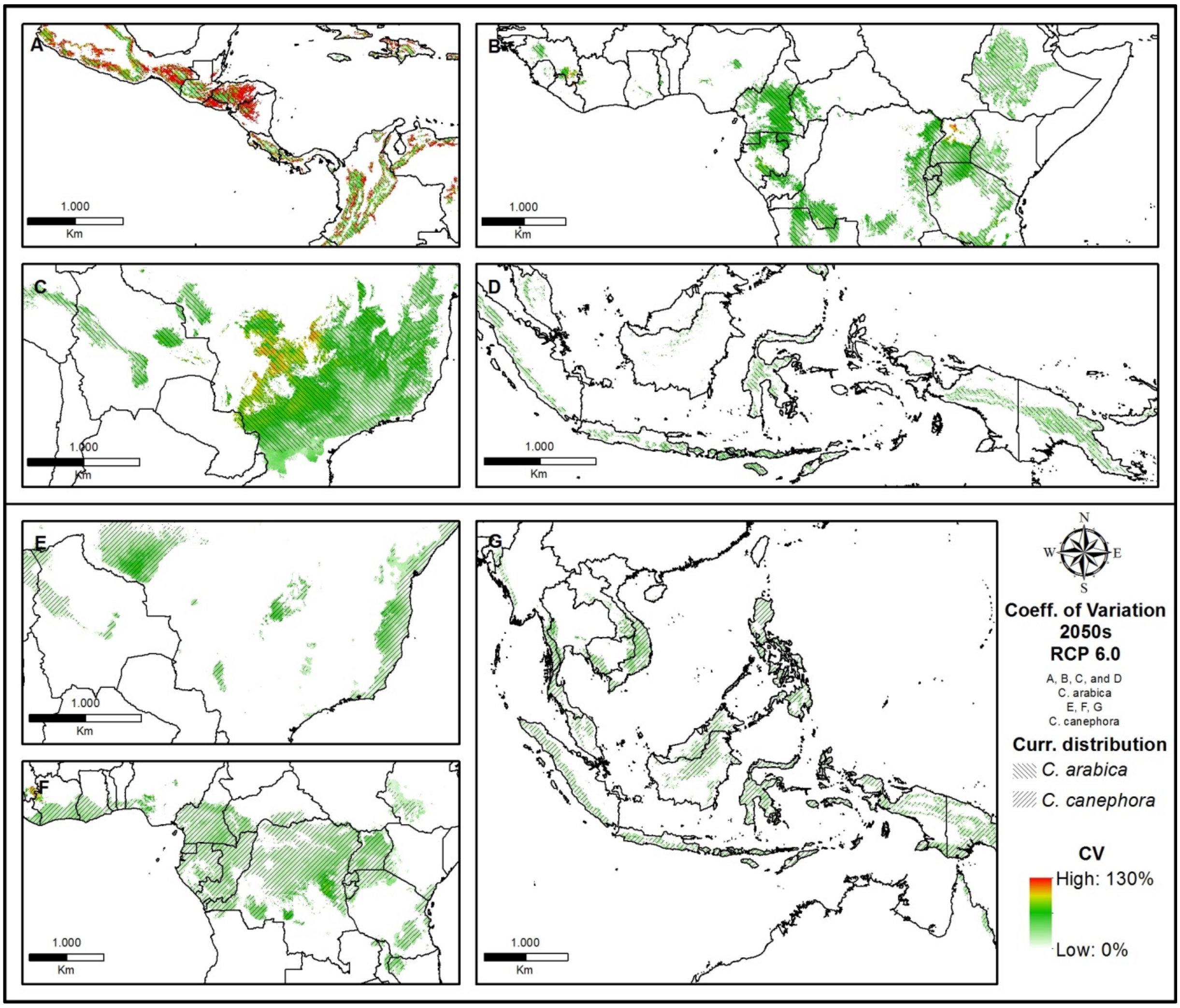
Figure 44. Suitability changes by the 2050s in the RCP 6.0 scenario



A-D: Arabica, E-G: Robusta; hatching indicates the current suitability distribution; warm colors represent areas with negative climate change impacts and cold colors positive changes (own data and representation).

The coefficient of variation across the 5 GCMs can be found in Figure 45. The CV was generally low, with the exception of Central America and the region around Brasilia in Brazil where it was up to 100% for *C. arabica*.

Figure 45. Coefficient of variation for 5 GCM outputs by the 2050s in the RCP 6.0 scenario



A-D: Arabica, E-G: Robusta; low coefficients of variation are shown as white, intermediate scores green and high model disagreement in red; hatching indicates the current suitability distribution (own data and representation).

3.2.3 AUC and Kappa to evaluate model accuracy

AUC values were consistently high across all model set ups. The lowest AUC value for Arabica coffee was .92 and .73 for Robusta, indicating that they performed much better than chance at discerning presence from background locations. Values for Cohen's Kappa were high on average but some models did not differentiate well between presence and background sample at an omission error threshold of 2.5% (Table 11).

Table 11. Average, minimum and maximum values of AUC, cAUC, Kappa and cKappa

	Average		Min		Max	
	Arabica	Robusta	Arabica	Robusta	Arabica	Robusta
AUC	0.97	0.93	0.92	0.73	1.00	1.00
cAUC	0.59	0.54	0.54	0.38	0.67	0.70
Kappa	0.31	0.21	0.02	0.00	0.98	0.99
cKappa	0.14	0.05	-0.16	-0.49	0.73	0.90

(Own data)

Considering the values that were compared to the performance of a simple null model, cAUC and cKappa, the majority of the machine learning models performed better than the distance based model. All models for Arabica coffee were better than the null model according to cAUC, and 77% according to cKappa. The Robusta models performed better than a null model in 74% of the cases according to cAUC and 50% according to cKappa.

ANOVA analysis of variance with the evaluation metric as dependent variable and the ensemble parameter levels as categorical independent predictor variables (Eq. 4) showed that these differences can be attributed to the chosen parameter variables. All independent variables (algorithm choice, background extent, background to presence point ration, and regularization level) contributed significant ($p < .01$) variation to the model evaluation statistics, AUC, cAUC, Kappa, and cKappa for both the Arabica and Robusta models.

3.2.4 Variable importance

Across all models for Arabica the mean temperature of the warmest quarter was ranked the most influential variable that contributed most to the suitability distribution. This was followed by the maximum temperature of the warmest month and mean temperature of the wettest quarter. The precipitation variables were ranked as least important, especially precipitation of the driest quarter and month (Bio 14 and 16). Among the temperature variables the two that indicated temperature variability (Bio 2 and Bio 7) were least influential.

The latter contrasted with the Robusta models where the mean diurnal range of temperature (Bio 2) and the annual temperature range (Bio 7) were consistently ranked high. This was followed by the maximum temperature of the warmest month. For Robusta the precipitation variables were given higher importance compared to the Arabica models. Among these, the precipitation variability (Bio 15) that described the intra-annual variation of precipitation was ranked highest. Least important were the temperature in the coldest quarter (Bio 11) and the precipitation during the coldest quarter (Bio 19) respectively (Table 12).

Table 12. Variable contribution as average variable weight in percent, and median rank of variable across all models

Bioclimatic variable		Average variable weight in %		Median rank of variable	
		Arabica	**Robusta**	**Arabica**	**Robusta**
BIO 1	Annual Mean Temperature	10	5	4	12
BIO 2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	4	12	12	1
BIO 3	Isothermality (BIO2/BIO7) (* 100)	4	2	9	15
BIO 4	Temperature Seasonality (standard deviati *100)	4	6	11	7
BIO 5	Max Temperature of Warmest Month	13	9	2	3
BIO 6	Min Temperature of Coldest Month	5	4	8	11
BIO 7	Temperature Annual Range (BIO5-BIO6)	2	11	16	2
BIO 8	Mean Temperature of Wettest Quarter	13	6	3	8
BIO 9	Mean Temperature of Driest Quarter	6	3	6	16
BIO 10	Mean Temperature of Warmest Quarter	15	7	1	7
BIO 11	Mean Temperature of Coldest Quarter	6	2	6	17
BIO 12	Annual Precipitation	2	5	15	8
BIO 13	Precipitation of Wettest Month	4	4	11	11
BIO 14	Precipitation of Driest Month	1	3	18	13
BIO 15	Precipitation Seasonality (Coefficient of Variation)	1	7	16	6
BIO 16	Precipitation of Wettest Quarter	3	4	13	12
BIO 17	Precipitation of Driest Quarter	1	4	18	10
BIO 18	Precipitation of Warmest Quarter	5	4	9	11
BIO 19	Precipitation of Coldest Quarter	2	2	17	15

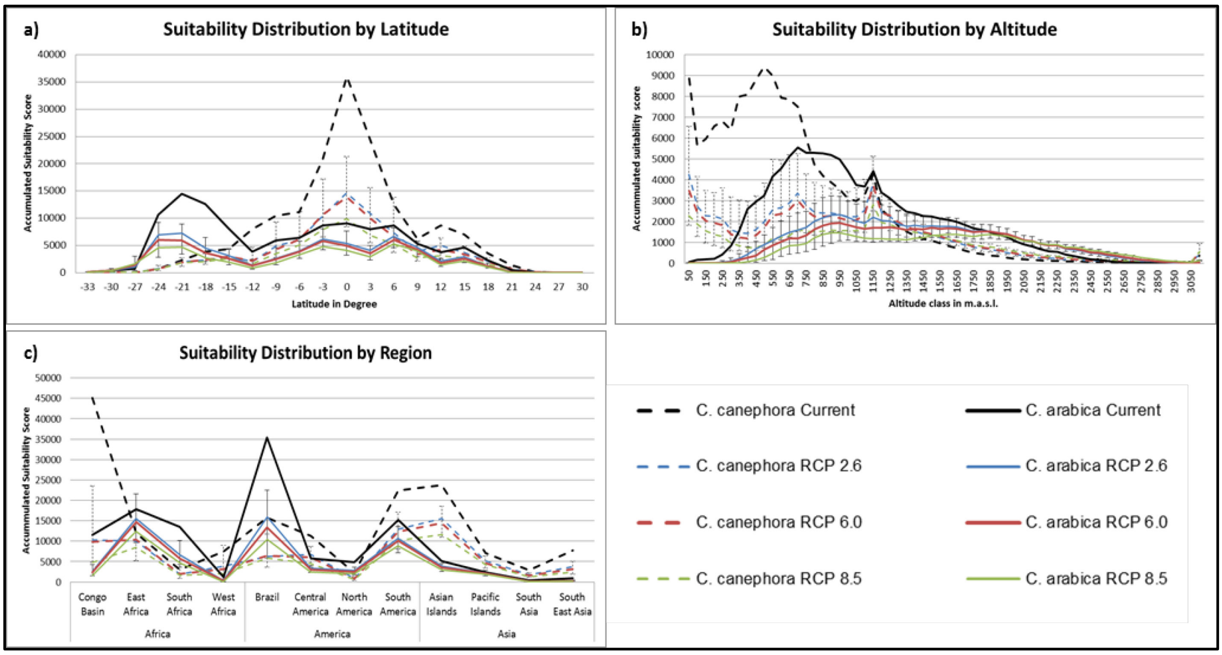
(Own data)

3.2.5 Distribution of climate change impacts

The suitability scores indicate how likely it is that a location is climatically suitable for coffee production. A higher sum of scores therefore means more suitable area. The sum of suitability scores across latitudinal meridians for current climate conditions and GCM outputs for scenario RCP 2.6, RCP 6.0 and RCP 8.5 is shown in Figure 46a indicating that *C. arabica* will lose suitability across all latitudes. Only in very high latitudes the losses were not as pronounced. Also for *C. canephora* losses occur mostly at low latitudes.

The sum of suitability scores in discrete altitude classes for both *C. arabica* and *C. canephora* by 2050 in the mean of RCP 2.6, RCP 6.0 and RCP 8.5 scenario GCM outputs and under current conditions is shown in Figure 46b. Both species will lose large shares of total suitability mostly in low altitudes below 1000 m.a.s.l. while relative losses in higher altitudes were not as drastic.

Figure 46. Distribution of aggregated suitability scores by latitude, altitude and region



a) Latitude, b) altitude, c) coffee regions; continuous lines represent *C. arabica*, dashed lines *C. canephora*, black lines the current distribution, colored lines future distribution; the error bars indicate the minimum and maximum across RCP 6.0 model means (own data and representation).

The sum of suitability scores for major coffee production regions for current conditions and of the mean of GCM outputs for the scenarios RCP 2.6, RCP 6.0 and RCP 8.5 by 2050 is shown in Figure 46c. The largest relative loss of suitability could be observed in Brazil and South East Asia for Arabica coffee. In these regions accumulated suitability score losses ranged between 85% in the RCP 8.5 scenario and 30% in the RCP 2.6 scenario. The least impact on Arabica was projected for East Africa and the Pacific Island region with 10% of suitability lost in the RCP 2.6 scenario and up to 30% in the RCP 8.5 scenario. Globally, losses were projected to be 49% of overall suitability score lost in the RCP 6.0 scenario (Table 13).

C. canephora suitability will be lost in the Congo basin. There, between 60% (RCP 2.6) to 95% (RCP 8.5) of total suitability may be lost in the species region of origin. Again, East Africa was projected with the least impact. In the RCP 2.6 scenario 16% of suitability will be lost here, and 30% in the RCP 8.5 scenario. Three of the important Robusta production regions, Brazil, South-East Asia and West Africa, were projected to experience losses of about

60% of suitability score. The global losses in our model were higher for Robusta (54%) than for Arabica. Even in the low impact scenario RCP 2.6 losses could be 51%. (Table 13).

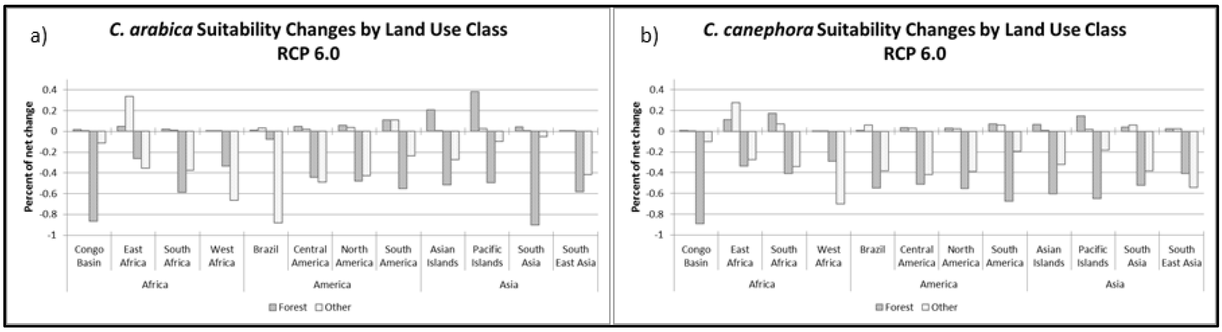
Table 13. Suitability by region in percent of current suitability; min/max values were based on the lowest/highest value found grid cells.

		C. arabica								
		RCP 2.6			**RCP 6.0**			**RCP 8.5**		
		Min	**Mean**	**Max**	**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**
Africa	Congo Basin	17	**24**	40	13	**20**	37	10	**15**	22
	East Africa	58	**87**	124	54	**82**	121	42	**70**	105
	South Africa	36	**49**	70	27	**43**	75	21	**34**	55
	West Africa	15	**28**	45	9	**19**	54	5	**10**	23
America	Brazil	24	**45**	71	16	**38**	64	11	**29**	48
	Central America	44	**59**	76	39	**52**	69	29	**42**	59
	North America	40	**57**	78	35	**52**	74	28	**43**	64
	South America	52	**71**	92	47	**66**	90	37	**58**	83
Asia	Asian Islands	55	**75**	98	50	**68**	88	38	**58**	78
	Pacific Islands	70	**91**	119	65	**86**	115	56	**79**	107
	South Asia	23	**40**	62	19	**33**	52	14	**27**	49
	South East Asia	24	**40**	63	18	**28**	46	12	**24**	49
Global		38	**57**	81	32	**51**	77	24	**42**	63
		C. canephora								
		RCP 2.6			**RCP 6.0**			**RCP 8.5**		
		Min	**Mean**	**Max**	**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**
Africa	Congo Basin	11	**23**	43	11	**22**	52	5	**11**	24
	East Africa	43	**84**	145	44	**87**	145	24	**70**	140
	South Africa	34	**63**	106	28	**64**	143	15	**53**	108
	West Africa	23	**52**	89	10	**42**	121	5	**27**	57
America	Brazil	26	**40**	82	23	**41**	75	19	**37**	75
	Central America	34	**60**	80	32	**54**	77	25	**40**	59
	North America	23	**41**	70	18	**33**	64	14	**28**	54
	South America	40	**58**	84	36	**55**	76	28	**45**	65
Asia	Asian Islands	50	**66**	88	46	**61**	78	35	**49**	67
	Pacific Islands	52	**74**	102	47	**66**	90	43	**60**	83
	South Asia	43	**61**	89	40	**56**	79	31	**43**	69
	South East Asia	31	**48**	72	24	**40**	65	17	**31**	60
Global		31	**49**	78	28	**46**	78	20	**36**	61

(Own data)

In Figure 47 the changes in suitability were distributed according to land use classes in the coffee producing regions by 2050 in the RCP 6.0 scenario. Globally, losses and gains in suitability were nearly equally distributed across area with forest cover and without forest cover. However, novel area made up for only about 10% of lost suitability for both species. For *C. arabica* exceptions are Brazil, East Africa and the Asian islands. In Brazil 90% of suitability losses were observed for areas without forest cover. In East Africa all of the suitable area lost that was currently not forest covered may be replaced with novel area that was not forest covered either. However, in Asia nearly all suitability gains were in areas with forest cover.

Figure 47. Distribution of suitability changes by region and land use classes by 2050 in RCP 6.0



Light grey represents area with forest cover and white bars area without forest cover; a) *C. arabica* b) *C. canephora* (own data and representation).

For *C. canephora* a similar pattern was observed. In West Africa 90% of suitability losses occur on land without forest cover, while in Asia Islands, Philippines and Indonesia, gains predominantly took place on land with forest cover. Also for Robusta most of the suitability losses in East Africa may be replaced by gains on open land. In the Congo basin large losses of suitability for both species will be observed on forested land.

3.3 Discussion

The goal of this chapter was to examine the implications of climate change for global coffee production. Analysis of changes in suitability under the RCP 6.0 scenario showed that climate change will reduce the area available for Arabica coffee in all world regions. Also Robusta will be less suitable in important regions in Brazil and Vietnam. Gains elsewhere will do little to offset these losses, giving global losses in suitability for both species of about 50 %. Only East Africa and the Asian island states showed substantial novel suitable area for both species to offset losses.

These results close a previously existing literature gap because this is the first study to model the impacts of climatic change on both coffee species globally using a reproducible approach. A novel methodology was developed here that was based on the notion that an ensemble of models captures more relevant information than a single model could. By using a mean of models on global scale based on several feasible parameter combinations rather the resulting analysis is more robust than previous regionally limited studies that used single models.

The extrapolation of the models with spatially explicit climate information resulted in global maps of suitability for both *Coffea arabica* and *Coffea canephora* based production which indicate suitability scores in major production regions. The underlying models were subsequently applied to the outputs of five global climate models for the RCP 2.6, RCP 6.0 and RCP 8.5 emission scenarios. The resulting maps were averaged for each emissions scenario and the change in suitability score was analyzed.

Both species show important changes in accumulated suitability scores at lower latitudes, which become less negative, albeit not positive, at higher latitudes. A southward latitudinal migration was also proposed by (Zullo et al. 2011) in a regional study in Brazil. However, no such impacts of climate change were found in other regions. Moreover, the gains in suitability in southern Brazil may not be enough to compensate for losses in suitability over large areas elsewhere. Similarly, losses in suitability are mostly at low altitudes while higher altitudes gain in suitability. (Schroth et al. 2009) and (Simonett 1988) identified similar altitudinal migration for Arabica in Central America and for Robusta in Uganda, respectively. These local studies confirm this analysis, which shows that altitudinal migration of coffee production will likely be a global trend. The magnitude of this effect, however, depends on how climate change will impact local conditions.

It had previously been hypothesized that Robusta production may be able to replace in part the losses in Arabica production. The suggestion was that *C. arabica* is heat sensitive and would thus suffer in a hotter world. In turn *C. canephora* may tolerate higher temperatures and could thus replace increasingly heat stressed Arabica coffee. While such a measure may be viable in some regions, our models draw attention to *C. canephora*'s need for climates with little intra-seasonal variability. As climate may not only become hotter, but also more variable, this may negatively affect Robusta coffee production. Thus, globally both species will be equally affected by climate change. The finding that the Congo basin, the origin of the species, may become almost entirely unsuitable by 2050 in a scenario with high emissions deserves further investigation as indigenous varieties are generally seen as the key to climate change adaptation of coffee.

An important factor in coffee production, especially for Robusta, is irrigation. This study included bioclimatic variables for natural rainfall distribution. Under current conditions most coffee production is rain-fed. But in the future farmers may increasingly turn to irrigation as a means of adaptation. An inclusion of meaningful irrigation information in future studies could thus alter some of our findings as such practices in the past have expanded the environmental range of production.

Analysis of suitability changes in several important coffee production regions for the RCP 6.0 emission scenario showed that for Arabica coffee production especially Brazil will see harmful climatic changes. Also Robusta may strive less well in important regions in Brazil and Vietnam. These losses were little counterbalanced with gains in other regions. Global losses amount to approximately 50% of lost suitability in the intermediate emissions scenario for both species. East Africa and the Asian island states appear to be the only regions with substantial gains in suitability for the two species. This finding partially contrasts with a proposed substantial reduction in climatically suitable area for indigenous Arabica varieties in Eastern Africa by (Davis et al. 2012). This difference suggests that commercial production has been adapted to a broader range of climatic conditions than can be found in Arabica's native range, but also highlights the need to use appropriate modeling parameters. However, given the long lifespan of coffee plantations the feasibility of a migration of coffee cultivation practices as proposed here needs to be investigated.

Interestingly, the East African areas with positive suitability changes are currently not covered with forest. In contrast, the Asian areas that gain suitability are currently under forest. The climate induced migration may thus result in further emissions from land use change. Whether or not newly suitable areas will be threatened by conversion from natural areas to agricultural land depends on economic incentives. This model showed that the currently highly productive regions of coffee production in Brazil and Vietnam may in the future become unsuitable. World markets through price effects may thus create such incentives. This would create economic opportunities in East Africa, but may induce additional deforestation in Asia, where Coffee is already a frontier crop. Policies designed to confront these challenges should thus be high on the agenda of stakeholders.

4 Where on earth is coffee grown? Spatial disaggregation of harvested area statistics using suitability data

Coffee production is of high significance for agricultural and conservation policies in its main production regions. Policies that address challenges from resource and land use conflicts are in high demand. Their evaluation requires an integrated systemic view of economic and physical spatial relationships as could be provided by a modeling framework like Globiom. However, to calibrate Globiom detailed data on the spatial distribution of production is needed.

Despite the demand for spatially explicit data its availability is still unsatisfactory especially in developing countries (Kuemmerle et al. 2013). Existing datasets of spatially disaggregated production statistics differ in their methodological approach and the resulting spatial distribution even for major crops like wheat (Anderson et al. 2015).

Coffee production shapes the landscape in its most important production regions (e.g. D'haeze et al. (2005) and plantations have become an integral part of ecosystems (Bosselmann 2012). Novel plantations, however, are often established at the frontier between forest and agricultural landscapes, driving deforestation (e.g. Hylander et al. (2013). Analysis of land use and land use change in these regions must therefore take coffee production into account.

Global trends also find repercussions in the coffee sector. Consumption is growing by approximately 2% a year (FAO 2014b) with large potential markets only starting to expand their demand (Lewin et al. 2004). Most of previous volume increases has come from increases in yields and less so from net area expansion (FAO 2014b). However, there has been a shift towards more competitive locations causing deforestation and resource depletion in novel origins (D'haeze et al. 2005).

Authorship:

The work presented in this chapter was conceived and carried out by myself. However, the cross-entropy approach for disaggregation in chapter 4.1.3 was proposed and implemented in GAMS by Aline Mosnier, International Institute of Applied Systems Analysis (IIASA) in Laxenburg, Austria.

Regional migration and increase deforestation pressure were also suggested by the results from chapter 3. Suitability losses will occur mostly at low elevations but novel suitability at high elevations, which are often not used for agricultural production (Bunn et al. 2015). However, only the suitability of area could be evaluated in the previous chapter. An evaluation of the impacts of these changes on global coffee production would have to relate these changes to actual coffee area harvested.

Available spatially explicit datasets of crop distribution are unsatisfactory in their representation of coffee production. (Monfreda, Ramankutty and Foley 2008) created a comprehensive set of spatially disaggregated crop production maps by combining agricultural census statistics with global cropland information derived from satellite data. This disaggregation based on political units did not take into account the climatic requirements of the coffee crop which are very specific. Furthermore, data is aggregated in a generic "coffee" category that spans both major coffee species. (You and Wood 2006) added the use of suitability maps based on climatic information to subnationally allocate coffee production. Nevertheless they too aggregated the data in a generic "green coffee" category[7].

The usefulness of these datasets is limited for the coffee industry and researchers specifically interested in coffee production (Eriyagama et al. 2014). The resource use and climatic requirements of the two main coffee species (*Coffea arabica* and *Coffea canephora*) are fundamentally different; an aggregation into a single category is therefore very limiting. Further differences in resource use exist between shaded agroforestry systems and sun grown coffee (van Rikxoort et al. 2014; Jha et al. 2014) and irrigated/rain-fed systems (Eriyagama et al. 2014).

As a step towards a better understanding of the spatial dynamics of coffee production associated land and resource use a method is described here to spatially disaggregate coffee harvested area data. The approach uses national production statistics, subnational land use statistics, satellite data of land cover, and climatic suitability information from chapter 3. The aim is to generate a globally spatially explicit dataset of coffee production for both of the two

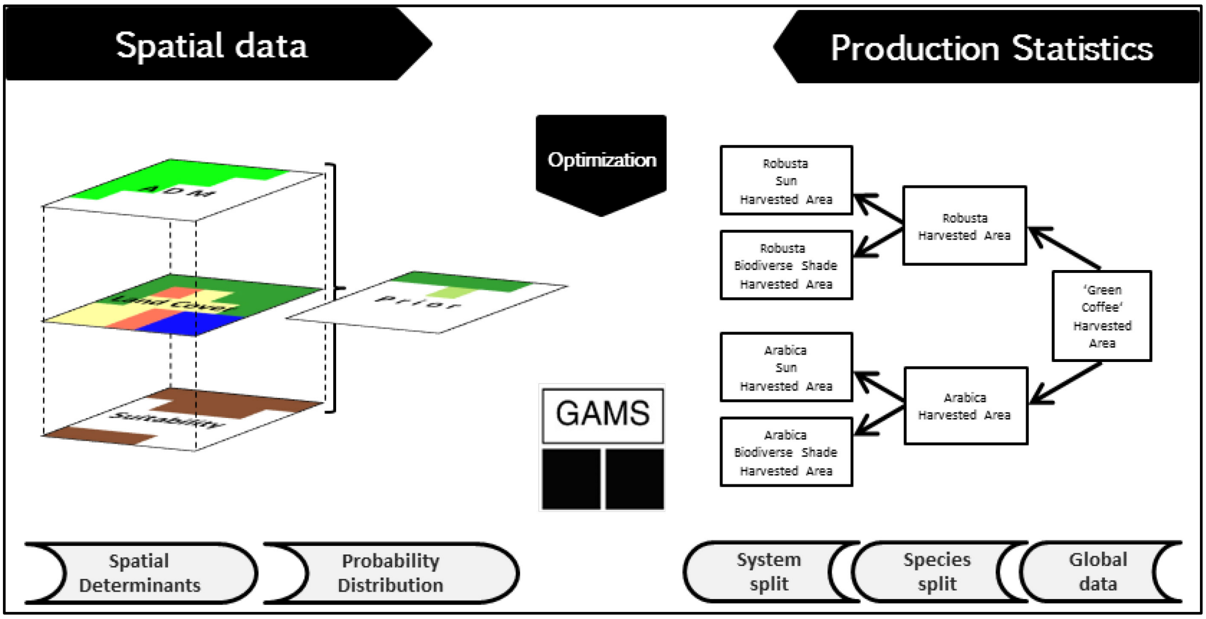
[7] Parallel to this work a novel MapSpam version was published that does differentiate between Arabica and Robusta (You et al. 2014). As the results for this chapter had already been produced, the new MapSpam dataset was not taken into account. It appears that the latest version is more useful for some regions, but nevertheless some data remains too unspecific. E.g. the Ivory Coast is classified as an Arabica producer even though it has historically been one the most important origins of Robusta style coffees.

main coffee species *C. arabica* and *C. canephora* that differentiates between sun grown production and biodiverse shade production.

4.1 Methods –Disaggregation of production statistics using crop suitability data as a prior probability

First, a database of national level harvested area statistics was disaggregated into data for production systems. Then, spatially explicit prior probabilities that a production system is present in a pixel cell were derived. Finally, the harvested area statistics database was spatially disaggregated using the prior probability (Figure 48). For evaluation independent data from subnational surveys were used as reference distributions.

Figure 48. Conceptual representation of the disaggregation approach



Spatial data was combined to derive a prior probability of presence for each production system. Aggregated national level statistics were split according to secondary data into production systems. This data was disaggregated by minimizing quadratic differences between area and prior using GAMS (own representation).

4.1.1 Data preparation

For total area harvested of both main coffee species data was used from the Food and Agriculture Organization of the United Nations (FAO) of the "green coffee" category. For coffee producing countries this database is the most complete and coherent database of production. This data was averaged over the years '98 – '02 (FAO 2012).

The next step was to separate the harvested area statistic into area for Arabica and area for Robusta. To do so production shares were calculated from data provided by the United States Department of Agriculture for the same time period (USDA 2012). The USDA data offered the most complete base for the time period in question of produced quantities separated by species. For Arabica it was assumed that 25% more area is needed than for Robusta because yields are generally considered to be lower for this species. For some countries the USDA data did not contain information about the species split. In such cases data from the International Coffee Organization database (ICO 2014) or internet resources were used to determine the main coffee species planted.

To disaggregate according to production systems data provided in a recent publication by (Jha et al. 2014) was employed. For this study the authors described the breakup of production systems of coffee according to shade and unshaded production based on expert opinion for most important coffee producing countries. These numbers were supplemented from additional data from the literature. Remaining gaps were filled by using the value of the nearest country with a known value.

The FAO harvested area statistics were thus disaggregated into Arabica/Robusta, and Shaded/Unshaded using the shares from USDA and (Jha et al. 2014). The result was a dataset of harvested area in 4 categories: Arabica - shaded, Arabica - unshaded, Robusta - shaded, Robusta - unshaded.

4.1.2 Prior probabilities

A prior probability of presence was calculated for each system based on three data sources: known subnational distributions, agro-climatic suitability and land cover class.

4.1.2.1 Subnational production data

Known subnational distributions of “green coffee” harvested area were downloaded from AgroMaps (FAO 2014a). This compiled survey data did not include continuous time series for all years. I therefore took data available for a period from 1990 to 2010, as close as possible to the year 2000 as a reference. Furthermore, the dataset did not offer data for all countries and the administrative resolution varied in detail. Rather than directly taking the area data from this dataset, for each pixel cell a probability between 0 and 1 was calculated based on the share of area within a reporting unit of total area. A country with unknown subnational distribution will thus have a homogenous equal probability across its entire area, while countries with known subnational distributions have 0 probabilities in spatial reporting units without production and a probability according to the share of total production in areas with production.

4.1.2.2 Coffee suitability data

For some countries subnational distribution was unknown; and for several others the known subnational distribution was rather coarse. Therefore the prior derived from the AgroMaps data was supplemented with climatic suitability maps for each species. From a set of 135 suitability maps for each species based on distinct algorithms and parameter settings the one was picked that showed the highest correlation with production data at different scales. As references a) the disaggregated harvest area statistics at national scale for Arabica and Robusta, b) subnational data from Philstat for Philippines for Arabica and Robusta (PSAI 2014), and c) subnational data from Brazil IBGE (2012) were used. The latter was aggregated for “green coffee”. In this case a multiple regression model was used with the two suitability models as predictors. The reasoning was that if the subnational distribution of suitability scores correlates with high resolution subnational distributions of harvested area then the suitability score is an appropriate means to estimate subnational distribution also in regions without detailed information.

The suitability maps from chapter 3 used different machine learning classification algorithms and various plausible parameter combinations (Bunn et al. 2015). Three algorithms were used to estimate the probability that a pixel cell is suitable for coffee production. GPS-referenced presence locations for each species and random samples from putative environments were employed as response variable. 19 bioclimatic variables downloaded from the WorldClim database were used as independent variables. The trained model was then applied to classify the corresponding set of interpolated climate data. Different parameters influenced the ability of the machine learning classification to correctly estimate distributions. Most influential were the ratio of random background samples to presence locations, the geographic extent from which the sample is drawn, and algorithm specific regularization parameters regulated the trade-off between model generalization and confidence.

The outputs of the classifications were continuous spatially explicit probabilities from 0 to 1 that a pixel cell’s climate is suitable for a species. Very low probabilities were assigned to large numbers of pixel cells, representing extreme conditions that are unlikely to be suitable. The suitability distribution for Arabica coffee was therefore truncated by applying a threshold based on a global 0% omission rate across the entire presence location sample. For countries with a presence point sample larger than 10 records a local 0% threshold was defined. For Robusta a 1% omission rate appeared more sensible on local level, and a 0% omission rate for the global threshold.

4.1.2.3 Land use data

The suitability score was only based on climate data and did not take into account land cover. The disaggregation should avoid assigning area to regions where current land use makes coffee cultivation unfeasible, e.g. urban areas, water bodies, or forest. The GLC2000 global land cover database offers a globally coherent classification of land cover for the year 2000 (European Commission 2003). The database comprises several classes of land cover, among these also cropland and other agricultural land. However simply limiting disaggregation to the cropland class would underestimate the true distribution of coffee production. The GLC2000 database used satellite data that is classified into the land cover classes. Agro-forestry systems like shaded coffee plantations are often difficult to discern from forest land based on satellite images (Tropek et al. 2014). Therefore, instead of limiting the disaggregation of coffee harvested area to the cropland category only, the land use classes were interpreted as probability of presence of a cropping system for coffee. Land use classes with “forest cover” were assigned a probability of 0.9 that coffee is produced under shade, and 0.1 that it is produced under sun. For the cropland land cover classes the probability for shade was 0.1 and 0.9 for intensive sun production. Mixed land use classes were assigned a probability of 0.5 each. Unfeasible land cover classes like bare areas or water bodies were given 0 probabilities (Table 14).

Table 14. **Global land use class definitions and production system probabilities**

	GLC Global Class (according to LCCS terminology)	Diverse Shade Probability	Sun Grown Coffee Probability
1	Tree Cover, broadleaved, evergreen LCCS	0.9	0.1
2	Tree Cover, broadleaved, deciduous, closed	0.9	0.1
3	Tree Cover, broadleaved, deciduous, open	0.9	0.1
4	Tree Cover, needle-leaved, evergreen	0.9	0.1
5	Tree Cover, needle-leaved, deciduous	0.9	0.1
6	Tree Cover, mixed leaf type	0.9	0.1
7	Tree Cover, regularly flooded, fresh water	0.9	0.1
8	Tree Cover, regularly flooded, saline water,	0.9	0.1
9	Mosaic: Tree cover / Other natural vegetation	0.5	0.5
10	Tree Cover, burnt	0.5	0.5
11	Shrub Cover, closed-open, evergreen	0.5	0.5
12	Shrub Cover, closed-open, deciduous	0.5	0.5
13	Herbaceous Cover, closed-open	0.5	0.5
14	Sparse Herbaceous or sparse Shrub Cover	0.5	0.5
15	Regularly flooded Shrub and/or Herbaceous Cover	0.5	0.5
16	Cultivated and managed areas	0.1	0.9
17	Mosaic: Cropland / Tree Cover / Other natural vegetation	0.1	0.9
18	Mosaic: Cropland / Shrub or Grass Cover	0.1	0.9
19	Bare Areas	0	0
20	Water Bodies (natural & artificial)	0	0
21	Snow and Ice (natural & artificial)	0	0
22	Artificial surfaces and associated areas	0	0
23	NoData	0	0

"Shade coffee"-categories are indicated in green, "sun-coffee" in red, mixed categories in blue, grey categories were excluded (own representation of European Commission 2003 and own data).

4.1.3 Disaggregation

The prior probability p in pixel i for each species j and system l was then given by (Eq. 5):

$$p_{ijl} = agromaps_i * suitability_{ij} * landcover_{il} \qquad \text{(Eq. 5)}$$

Where for each pixel $agromaps_i$ is the probability of presence of coffee production from known subnational distributions, $suitability_{ij}$ is the probability for each species based on the classification of climate, and $landcover_{il}$ is the probability for each system based on current land use.

Disaggregation of harvested area by species and system was done by minimizing equation 6 in GAMS 23.8.2:

$$\min \sum_i \sum_j \sum_l \left(\frac{Area_{ijl}}{\sum_{Country} Area_{jl}} - \frac{p_{ijl}}{\sum_{Country} p_{jl}} \right)^2 \forall country \qquad \text{(Eq. 6)}$$

s.t.

$$\sum_i Area_{ijl} = Area_{Country} \, \forall j \forall l \qquad \text{(Eq. 7)}$$

$$\sum_j \sum_l Area_{ijl} \leq Available\ Area_i \forall i \qquad \text{(Eq. 8)}$$

Conditions were that all area for each country by species and system is disaggregated to grid cells (Eq. 7), and that the area within a grid cell does not exceed available area in a grid cell (Eq. 8). Available area was determined by the highest concentrations of coffee production observed as follows: A maximum of 20% of land is used for coffee production to coffee in Vietnam (D'haeze et al. 2005). In Brazil the situation is similar where in Minas Gerais 10% of agricultural land is under coffee production (IBGE 2012). If a min of arc has 1852m at the equator, a 5arcmin pixel cell has $(5*1852)^2=85747600m^2$ or 8574.76ha. The maximum area was therefore estimated as 1715ha (20% of total grid cell area) under sun and 857ha (10% of total grid cell area) under shade.

4.1.4 Evaluation

Independent data with high spatial resolution as reference datasets to evaluate the disaggregation results was difficult to obtain. We compared the resulting maps for Arabica and Robusta production under shade and sun with (a) a database of known occurrence locations, (b) subnational survey data not included in AgroMaps, and (c) the Monfreda and MapSpam "green coffee" data.

The occurrence location data from chapter 3.1.1 was used to evaluate the accuracy of disaggregation. An occurrence location is a location of actual coffee production of either Arabica or Robusta coffee. Even though the database did not cover all global production regions at the occurrence sites coffee area should be (a) assigned and, (b) higher than within a random sample from coffee producing countries. For evaluation of disaggregation results the distribution of area at Arabica and Robusta occurrence sites was compared with the distribution of area in an equally sized random sample drawn. For comparison with previous disaggregation efforts that do not differentiate species all occurrence sites were pooled. The random sample was taken from the entire area of coffee producing countries, except from within a 25km buffer around occurrence sites.

Comparison with independent survey data was done by mapping the respective distributions. As reference coffee production survey data that was not included in the AgroMaps subnational database was used. Four data sets were available, although not always from the original source: Census data from Brazil, Costa Rica, Ethiopia and Vietnam (Table 15).

Table 15. Census data used for visual evaluation of downscaling results

Country	Coffee species in country	Survey data source	Survey year	Available reference
Brazil	Arabica/ Robusta (sun)	IBGE - Instituto Brasileiro de Geografia e Estatística	2006	IBGE 2012
Costa Rica	Arabica	INEC - Instituto Nacional de Estadística y Censos (Costa Rica)	2006	Instituto Nacional de Estadística y Censos 2006
Ethiopia	Arabica	IFPRI – International Food Policy Research Institute (Tadesse et al. 2006)	2006	Rüegsegger 2008
Vietnam	Arabica/ Robusta	VICOFA 2004	2004	Giovannucci et al. 2004

(Own data)

The analysis aimed at the comparison of the spatial patterns found in the survey data and the downscaled data. Ideally, the distribution of area in the survey data would be identical with the data that resulted after downscaling. A direct comparison of the data was not feasible due to the differing spatial aggregation.

Two kinds of comparisons were done. Survey data for Brazil and Vietnam discriminates between the two species. For these two countries visual comparison allowed the assessment of the accuracy of the downscaling discrimination by species. By pooling area from both species into a green coffee category a comparison with previous datasets that used this definition was possible. In Ethiopia and Costa Rica only one species (*C. arabica*) is used. The pooled data is therefore identical with the Arabica data, allowing the simultaneous evaluation of disaggregation accuracy and comparison with previous data.

4.2 Results – A spatially explicit database of coffee production

First the preparation of data will be shown: split of area by production system, and the choice of a suitability map for prior preparation. Then, the results of the disaggregation step will be compared with subnational reference data, known area distribution, and previous disaggregation efforts.

4.2.1 Hectares by production systems

Table 16 shows the split of FAO "green coffee" harvested area data into the four production systems Arabica-sun, Arabica-shade, Robusta-sun, and Robusta-shade.

Table 16. Split of FAO "green coffee" harvested area data into four production systems

Country	FAO area harvested in ha	Arabica area in ha	Robusta area in ha	Share of biodiv. shade %	Arabica area shade in ha	Arabica area sun in ha	Robusta area shade in ha	Robusta area sun in ha
Angola	27600	0	27600	13	0	0	3627	23973
Belize	51	51	0	34	17	34	0	0
Benin	950	0	950	75	0	0	713	238
Bolivia	24362	24362	0	38	9136	15226	0	0
Brazil	2253650	1928381	325270	0	0	1928381	0	325269
Burundi	27800	27749	51	0	0	27749	0	51
Cambodia	365	0	365	5	0	0	20	345
Cameroon	223525	22529	200996	30	6759	15770	60299	140697
Central African Rep.	15982	0	15982	34	0	0	5458	10525
China	10240	10240	0	36	3650	6590	0	0
Colombia	737140	737140	0	15	110571	626569	0	0
Comoros	620	0	620	0	0	0	0	620
Costa Rica	108903	108903	0	5	5445	103458	0	0
Côte d'Ivoire	717818	0	717818	75	0	0	538364	179455
Cuba	60200	60200	0	68	41164	19036	0	0
D. R. Congo	116540	16419	100120	22	3641	12778	22204	77917
Dom. Rep.	138166	138166	0	95	131258	6908	0	0
Ecuador	314649	189912	124737	80	151930	37983	99790	24947
El Salvador	161883	161883	0	58	93892	67991	0	0
Eq. Guinea	10426	0	10426	30	0	0	3128	7298
Ethiopia	239922	239922	0	95	227926	11996	0	0
Fiji	25	0	25	25	0	0	6	19
French Polynesia	88	0	88	20	0	0	18	70
Gabon	408	0	408	30	0	0	122	286
Ghana	12798	0	12798	75	0	0	9598	3199
Guadeloupe	33	0	33	68	0	0	22	10
Guatemala	264800	263501	1298	42	110671	152831	545	753

Guinea	49359	0	49359	75	0	0	37020	12340
Guyana	276	276	0	0	0	276	0	0
Haiti	55949	55949	0	95	53152	2797	0	0
Honduras	207281	207281	0	25	51820	155461	0	0
India	306198	130682	175516	65	84944	45739	114085	61430
Indonesia	1201588	102800	1098788	25	25700	77100	274697	824091
Jamaica	5080	5080	0	95	4826	254	0	0
Kenya	171700	171477	222	15	25722	145755	33	190
Laos	31227	0	31227	6	0	0	1720	29507
Liberia	14015	0	14015	75	0	0	10511	3504
Madagascar	193227	15707	177520	0	0	15707	0	177520
Malawi	3292	3292	0	0	0	3292	0	0
Malaysia	48000	0	48000	21	0	0	10032	37968
Mexico	715055	695258	19796	20	139052	556207	3959	15837
Mozambique	1122	0	1122	0	0	0	0	1122
Myanmar	4310	0	4310	34	0	0	1482	2828
Nepal	376	376	0	65	245	132	0	0
New Caledonia	270	270	0	25	68	203	0	0
Nicaragua	103246	103246	0	40	41299	61948	0	0
Nigeria	3463	0	3463	33	0	0	1147	2316
Panama	27413	27413	0	30	8224	19189	0	0
Papua New Guinea	81694	78798	2896	25	19700	59099	724	2172
Paraguay	4874	4874	0	0	0	4875	0	0
Peru	245399	245399	0	90	220859	24540	0	0
Philippines	134358	8728	125630	20	1787	6942	25716	99914
Puerto Rico	30482	30482	0	95	28958	1524	0	0
Rep. Congo	4754	0	4754	30	0	0	1426	3328
Rwanda	24945	24945	0	18	4450	20496	0	0
Samoa	34	0	34	20	0	0	7	28
Sierra Leone	9949	0	9949	75	0	0	7462	2487
Sri Lanka	15614	4258	11355	65	2768	1490	7381	3975
Suriname	211	0	211	0	0	0	0	211
Tanzania	112754	86979	25774	0	0	86980	0	25775
Thailand	67418	217	67201	6	13	204	4032	63169
Togo	48378	0	48378	75	0	0	36284	12095
Trinidad and Tobago	2640	207	2433	0	0	207	0	2433
Uganda	264499	40160	224339	55	22088	18072	123386	100953
United States	2550	2550	0	23	599	1952	0	0
Vanuatu	42	0	42	25	0	0	10	31
Venezuela	224079	224079	0	9	19844	204235	0	0
Vietnam	386200	3900	382299	5	195	3706	19115	363184
Yemen	33099	33099	0	95	31445	1655	0	0
Zambia	4980	4980	0	0	0	4980	0	0
Zimbabwe	6249	6249	0	0	0	6249	0	0
SUM	10,316,603	6,248,379	4,068,223		1,683,814	4,564,566	1,424,145	2,644,079

(Own data compiled from USDA 2012; ICO 2014; FAO 2012; Jha et al. 2014)

4.2.2 Choice of suitability map

For each species a total of 135 suitability maps from chapter 3 (Bunn et al. 2015) were tested that differed in the algorithm and parameter settings used. The final suitability maps with the highest correlation coefficients were based on the probabilistic output of the RandomForest classification algorithm. The background sample for this model used a 5:1 background to presence sample ratio. The sample was drawn from all locations that were within the range of annual mean temperature found in the presence sample. As regularization 1000 trees were constructed with 8 variables picked at each node. The suitability maps obtained this way showed significant correlation with actual production (Table 17).

Table 17. Correlation coefficients of suitability maps with reference distributions

Dataset	Correlation coefficient	df	p-Value
National Arabica	.897	69	.000
National Robusta	.623	69	.000
Brazil "green coffee"	.259*	5486	.000
Philstat arabica	.170	77	.133
Philstat robusta	.401	77	.000

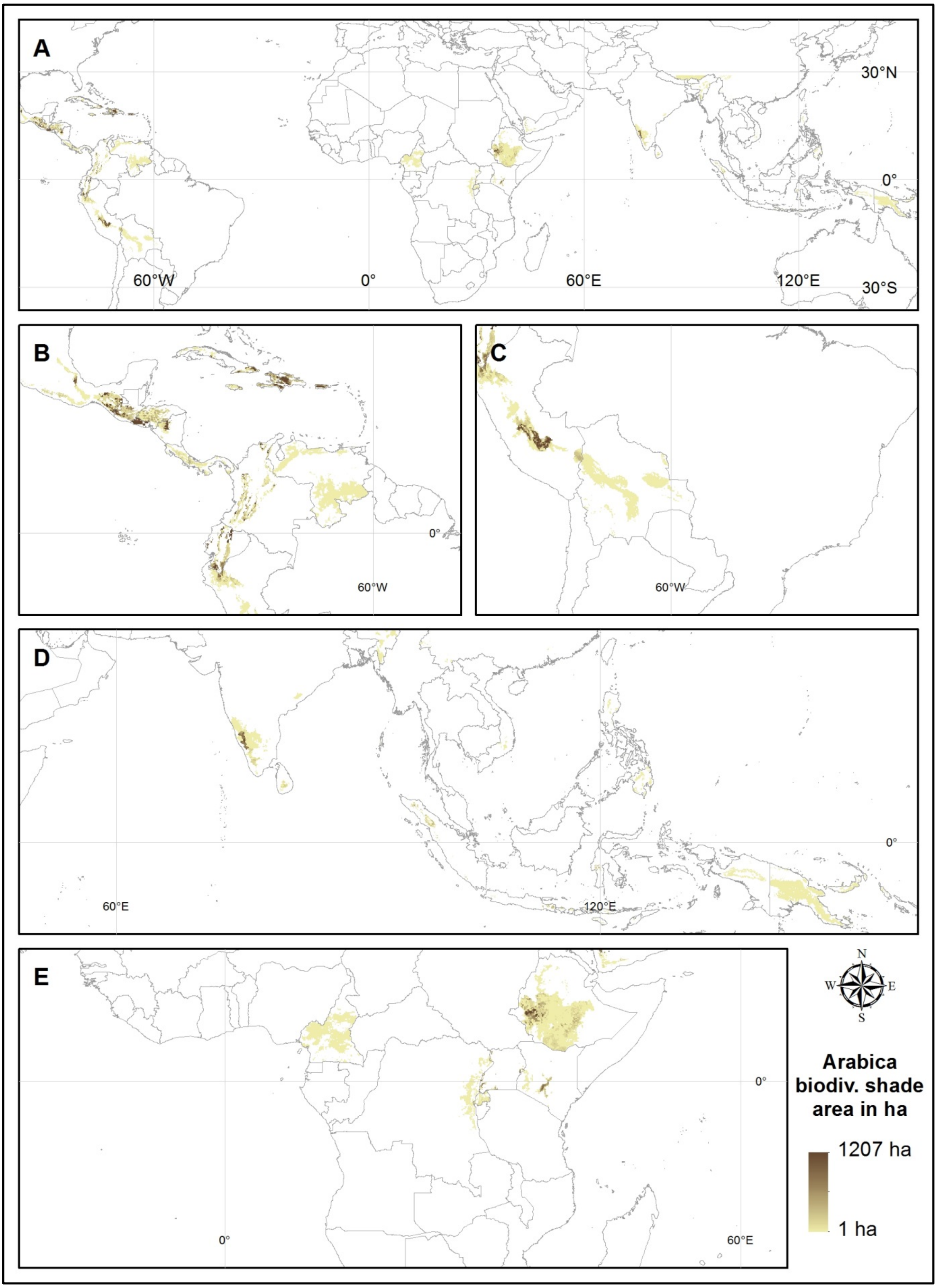
(Own data) *R^2

On national level the Arabica suitability scores were highly correlated with harvested area of Arabica. In the case of Brazil the R^2 showed that the suitability maps reflected to some extent the subnational distribution of coffee production with both suitability maps significantly adding to the model (data not shown). However, correlation of Arabica suitability with Arabica harvested area as reported in PhilStat was not significant with high confidence.

4.2.3 Maps of disaggregated data

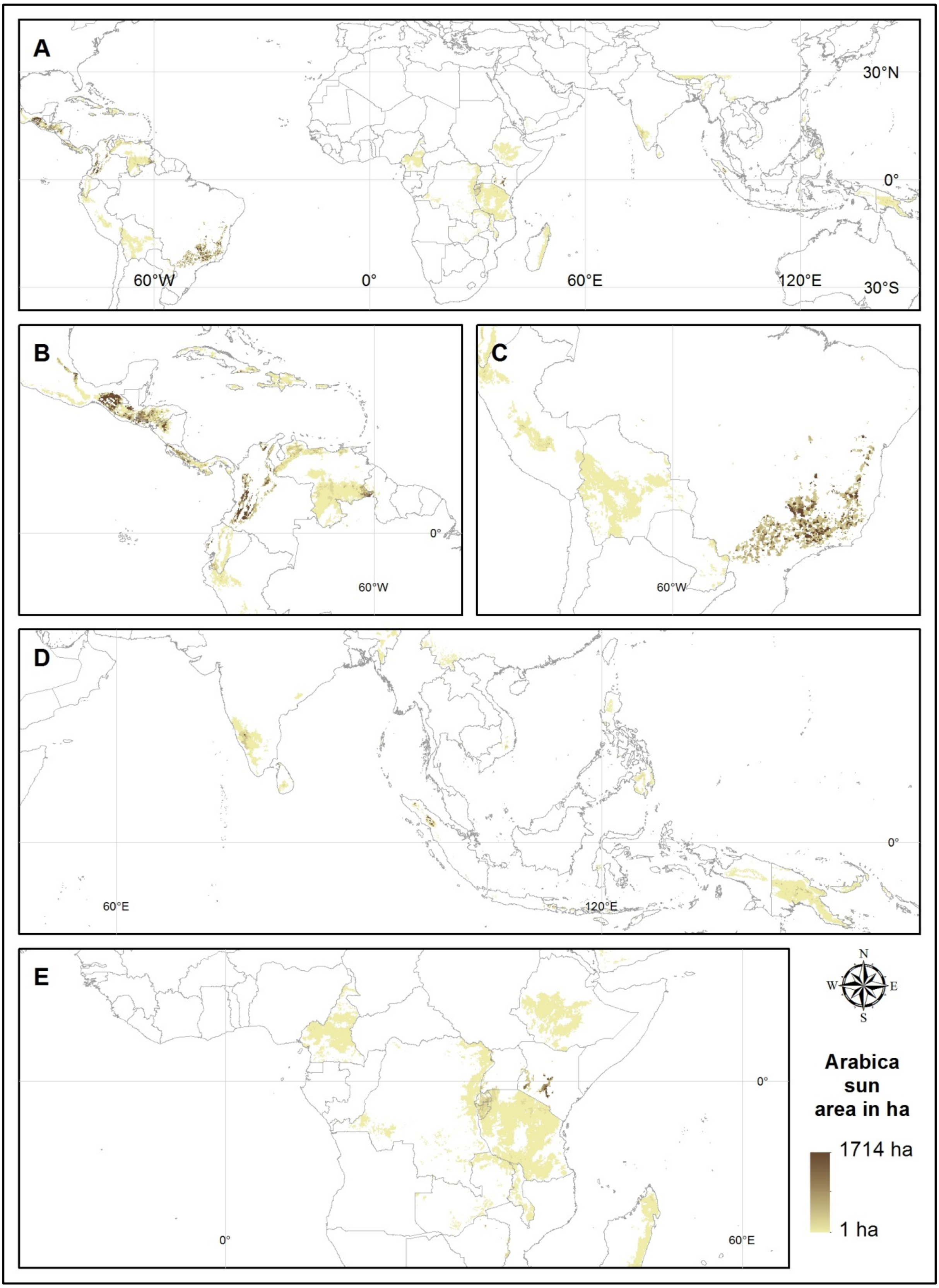
Figure 49-Figure 52 show the distribution of area for coffee production in the four production systems. Robusta based production is much more spatially dispersed than Arabica production. Most Robusta coffee is produced under full sun. Ivory Coast, India, Indonesia and Ecuador were the countries with significant Robusta production in biodiverse shade. While Robusta is grown in large regions in Africa and Asia, Arabica production is much more concentrated in South America.

Figure 49. Global distribution of Arabica area under biodiverse shade in ha



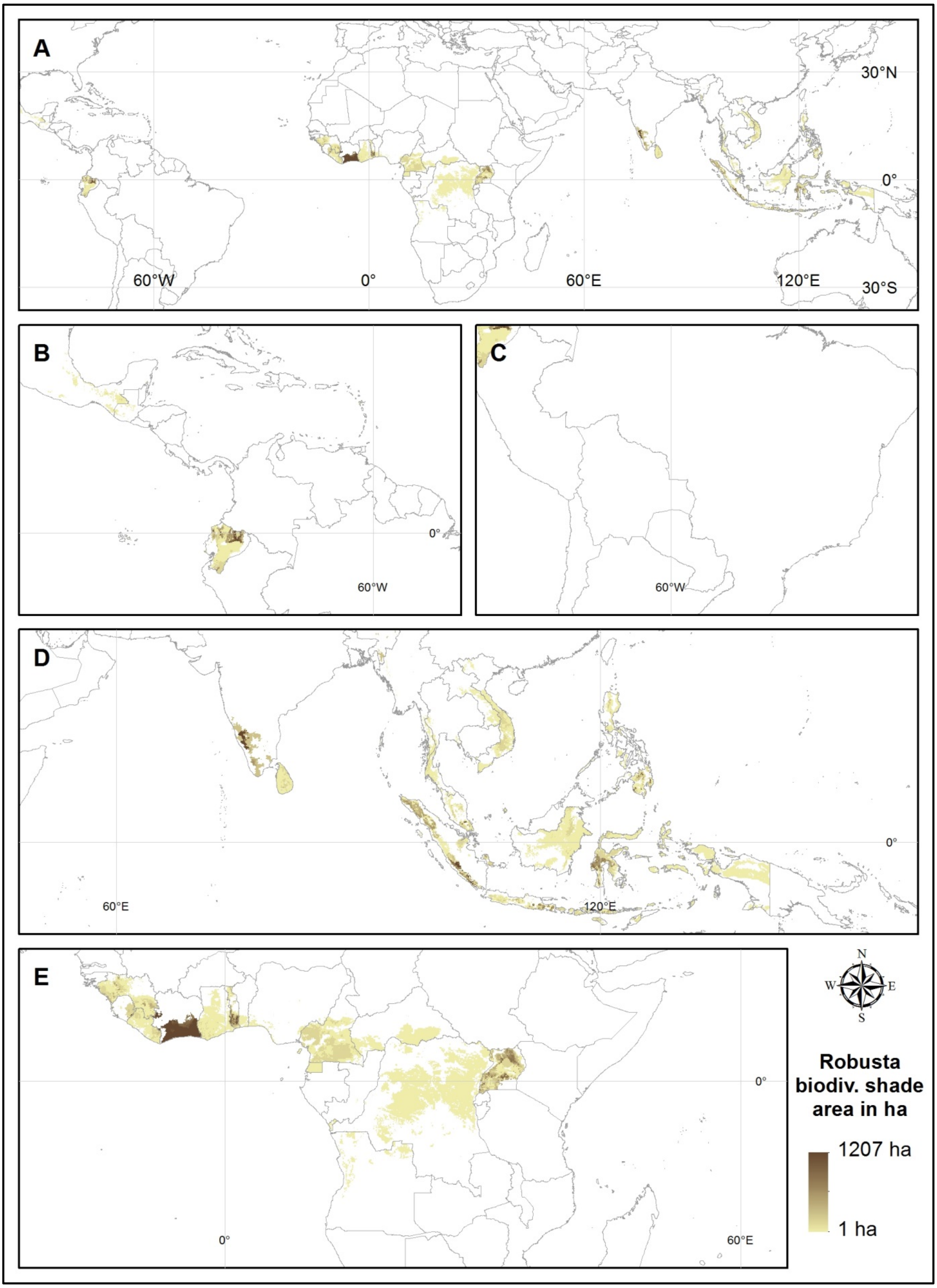
Ha per grid cell; A) Global; B) Central America; C) Brazil; D) South Asia; E) Africa (own data and representation).

Figure 50. Global distribution of Arabica area under sun



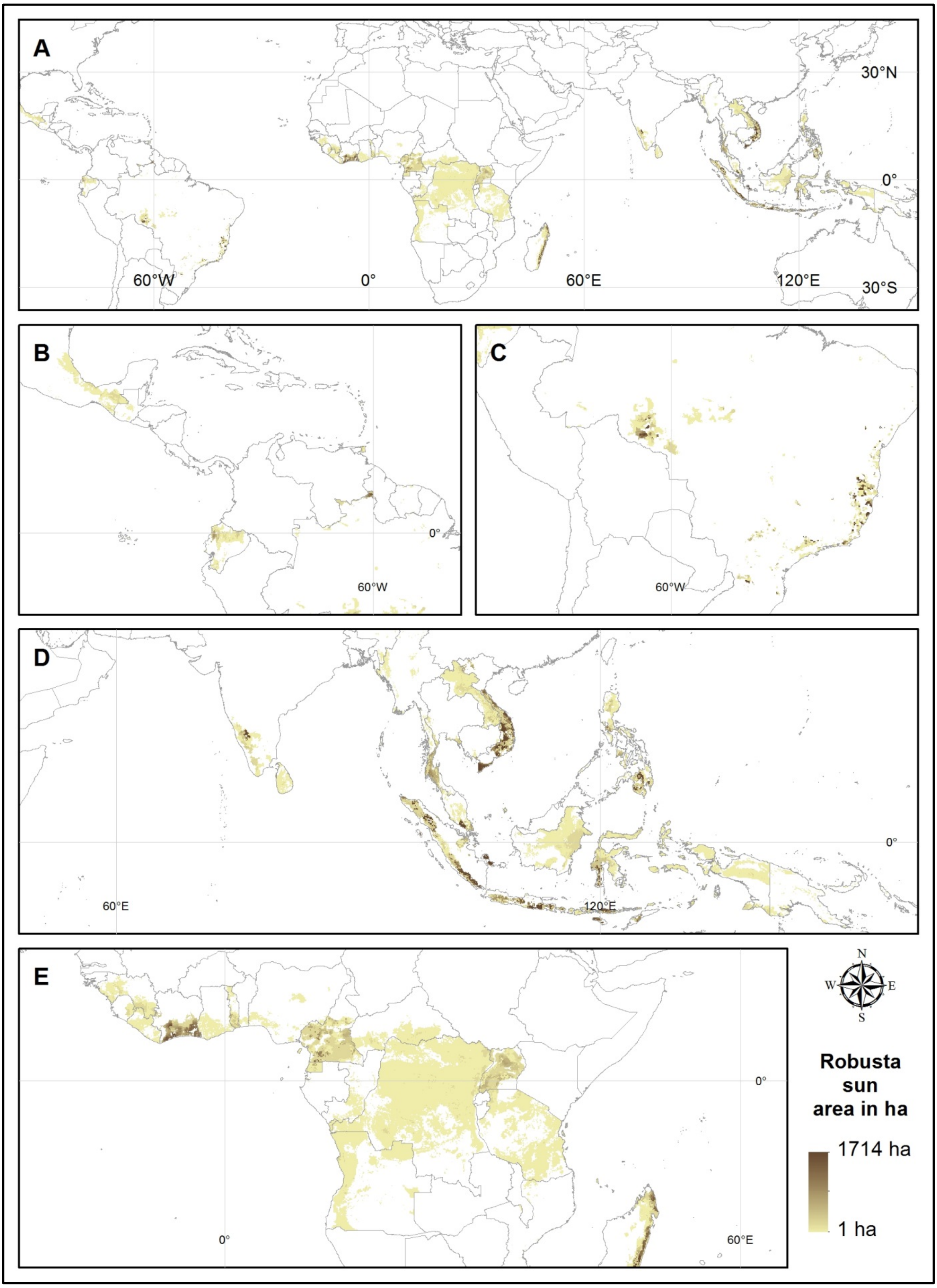
Ha per grid cell; A) Global; B) Central America; C) Brazil; D) South Asia; E) Africa (own data and representation).

Figure 51. Global distribution of Robusta area under biodiverse shade in ha



Ha per grid cell; A) Global; B) Central America; C) Brazil; D) South Asia; E) Africa (own data and representation).

Figure 52. Global distribution of Robusta area under sun in ha

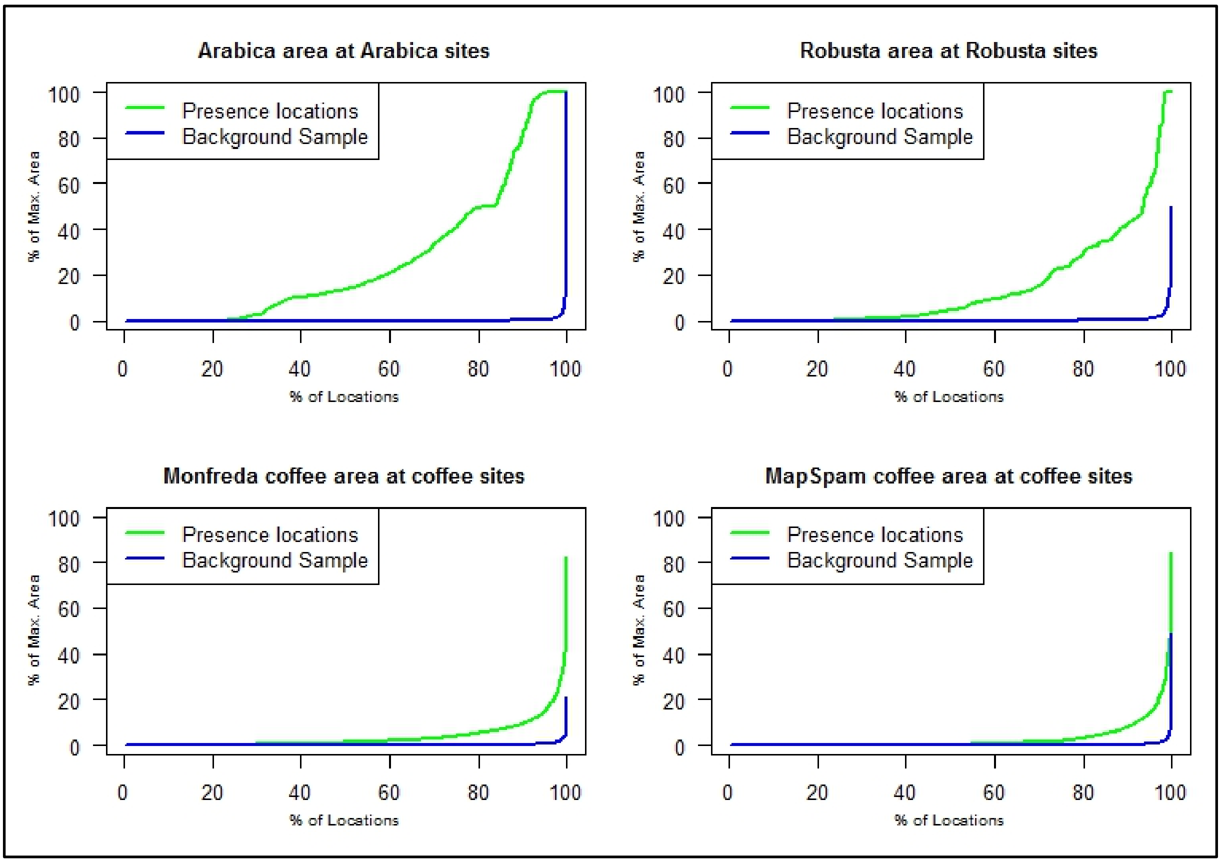


Ha per grid cell; A) Global; B) Central America; C) Brazil; D) South Asia; E) Africa (own data and representation).

4.2.4 Area at occurrence locations

The distribution of area at Arabica and Robusta occurrence sites was compared with the distribution of area in an equally sized random sample for both this disaggregation effort, as well as for MapSpam and Monfreda (Figure 53). Graphs show percentages relative of the maximum area for comparison. At Arabica locations about 20% of occurrence locations and 85% of background locations were assigned marginal to no area. Approximately 10% of locations were near the available area per pixel cell area limit, which can also be observed at some background locations (Figure 53 A). A similar percentage of Robusta occurrence locations was omitted in the model. The following third of locations was assigned relatively less area than at Arabica locations. Only few locations were limited by the area constraint. No background locations were assigned large area but the commission error appeared higher than for the Arabica model as more background locations had marginal area (Figure 53 B).

Figure 53. Distribution of area at occurrence locations



For comparison percentage of maximum area is shown. A) Total Arabica area *at C. arabica* locations; B) Total Robusta area at *C. canephora* locations; C) Coffee area at coffee locations (Monfreda); D) Coffee area at coffee locations (MapSpam); Green lines show the perc. area at occurrence locations and blue lines at random background locations (own data and representation).

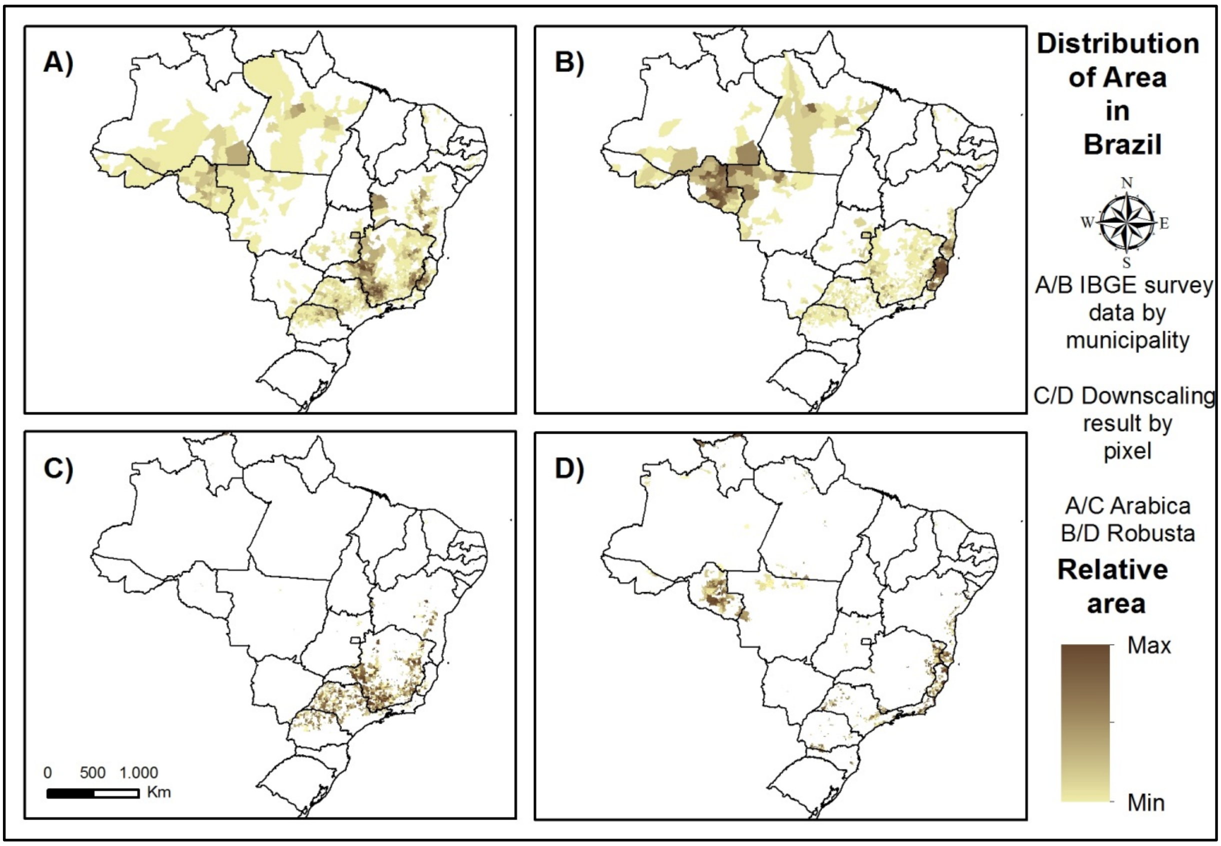
In comparison with the disaggregation by Monfreda and MapSpam the omission was reduced as shown by the relatively larger area at presence sites. 30% of presence sites were not assigned area in the Monfreda model, and about 50% in the MapSpam model. Additionally, the curves rise slower for the Monfreda and MapSpam model. Thus, most area is concentrated in a few pixel cells while most presence sites were assigned marginal values only. However, this was also the case for the background sample where only few sample sites were assigned area in the Monfreda and MapSpam models (Figure 53 C/D).

The method proposed here resulted in a more accurate assignation of area to occurrence locations of coffee than previous efforts. However, also more area was assigned to background locations which putatively do not produce coffee.

4.2.5 Comparison with subnational reference data

The census data distribution in administrative units of Brazil and Vietnam of Arabica and Robusta production was compared with the modeled distribution. In Brazil the distribution of Arabica production was reflected well in the model for the traditional coffee regions in the South, but production in non-traditional regions to the West of the country was underestimated (Figure 54 A/C). For Robusta main regions in the West and in coastal Espirito Santo were assigned area in the model but more central locations in Minas Gerais for example were omitted (Figure 54 B/D).

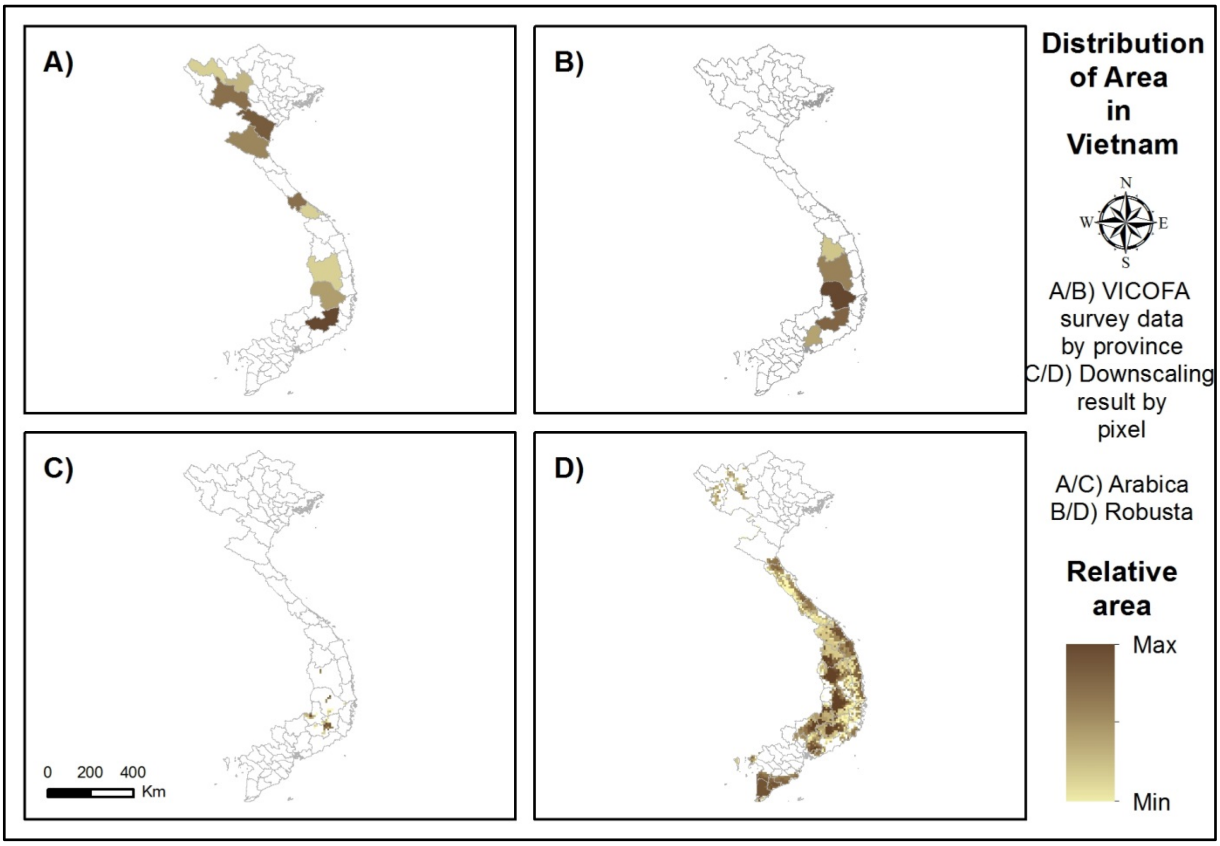
Figure 54. Distribution of area in Brazil for Arabica and Robusta production



Dark brown represents the maximum area, light brown the minimum; A/B show the distribution according to IBGE (2012) by municipality; C/D the modeled distribution by pixel; A/C is Arabica, B/D Robusta under sun (own representation).

For Vietnam the distribution of coffee production in the model somewhat alligned with the distribution acording to (VICOFA 2004). The most important production regions for both species were reflected in the model. However, the distribution of Robusta production area was overestimated both to the North and South of the key production region (Figure 55 A/C). Distribution of Arabica area on the other hand was underestimated. Production in locations in Northern Vietnam was omitted (Figure 55 B/D).

Figure 55. Distribution of area in Vietnam for Arabica and Robusta production



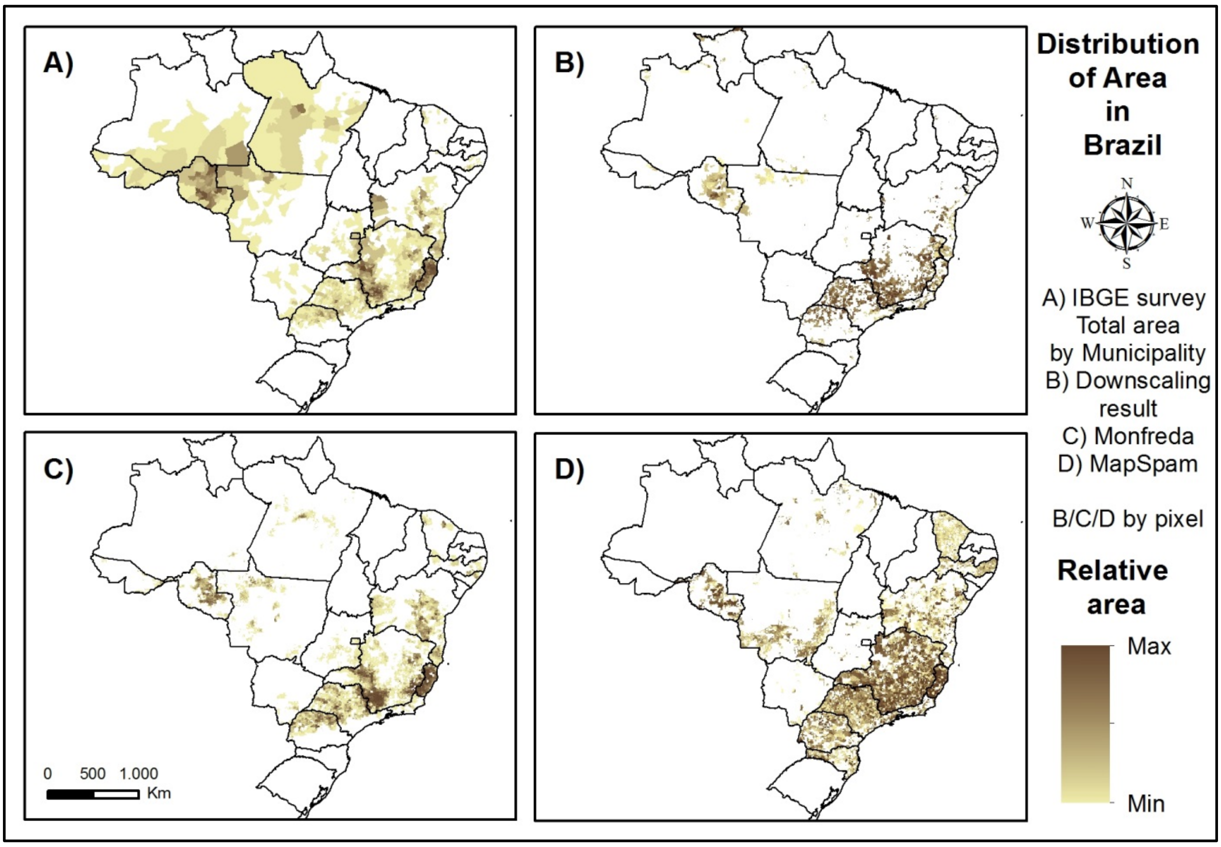
Dark brown represents the maximum area, light brown the minimum; A/B is the survey distribution (VICOFA 2004) by province; C/D the modeled distribution by pixel (A/C) Arabica data; (B/D) Robusta data (own representation).

4.2.6 Comparison with other efforts

Comparison of this disaggregation model with previous efforts showed the differences between the approaches. For comparison all data was aggregated into a single “green coffee” category because the results of Monfreda and MapSpam used this definition. Additionally, spatial aggregation of data and data format varies between sources so that the maps for evaluation show relative area rather than absolute values. Survey data for four countries was available: Brazil, Vietnam, Ethiopia and Costa Rica. The four countries differ in their production and data characteristics. Therefore, in the following the results for each country are discussed in the context of these characteristics.

Brazil is dominated by full sun production of Arabica, mixed with Robusta. Subnational data was available in AgroMaps. The association of coffee area with cropland type land cover as done by Monfreda and MapSpam is likely to be more correct with sun production than for biodiverse shade. All three efforts thus did not differ much in the disaggregation result. The spatial pattern found in the survey data was very similar to the one resulting from this disaggregation effort and the Monfreda data in main production areas. MapSpam appeared to overpredict some areas in the central coffee region. My approach underestimated areas in Northern states (Figure 56).

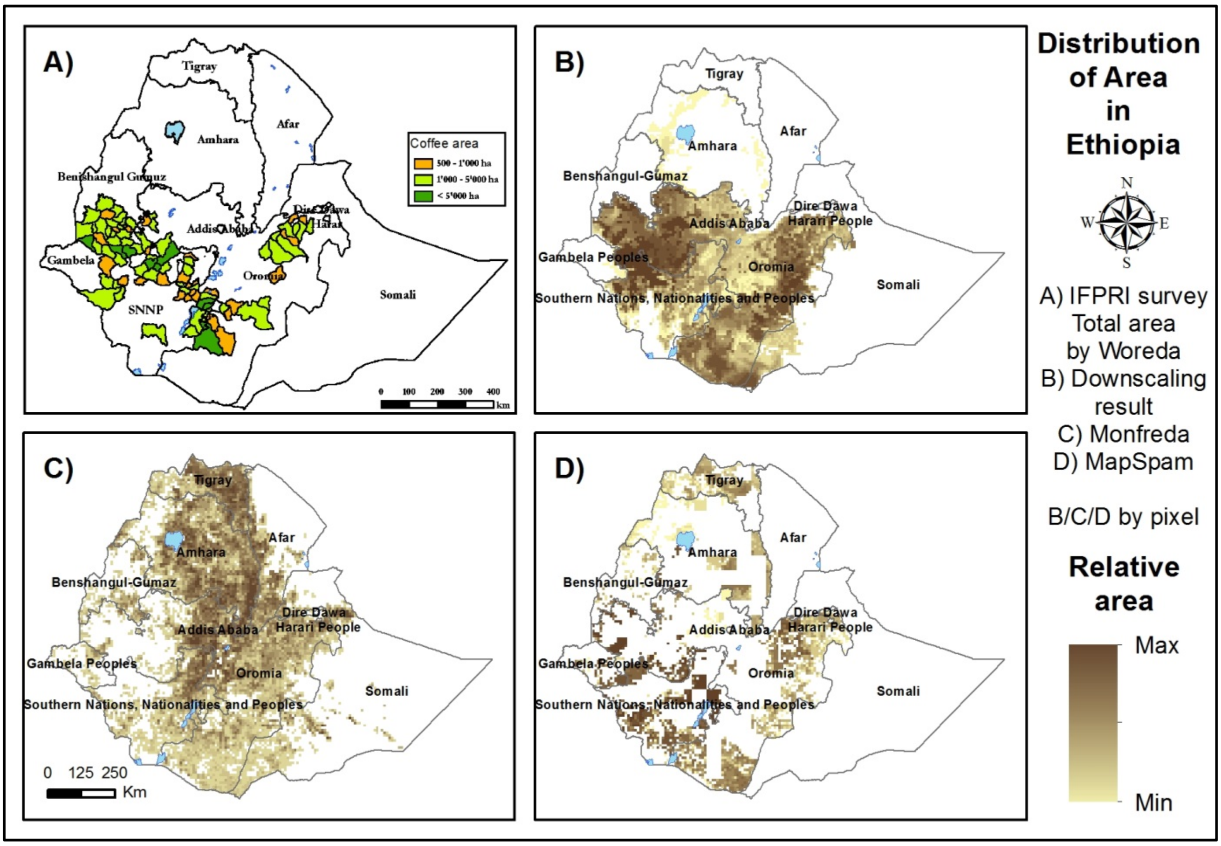
Figure 56. Comparison of area distribution of different data sources (Brazil)



Dark brown represents the maximum area, light brown the minimum; A) IBGE 2006 survey data; B) Total area aggregated over Arabica and Robusta; C) Monfreda; D) MapSpam (own representation).

Ethiopia does not produce Robusta style coffee. Most production is Arabica under biodiverse shade. No subnational data was available from AgroMaps. The disaggregation results differed substantially for this country. Compared to the survey data from IFPRI 2006 (Tadesse et al. 2006) adapted from (Rüegsegger 2008) the Monfreda land cover based effort appeared to place coffee outside of actual coffee regions. MapSpam reflected the actual pattern a little closer. The suitability map downscaling effort presented here described major coffee regions somewhat better but nevertheless overpredicted area in the Southern part of Ethiopia (Figure 57).

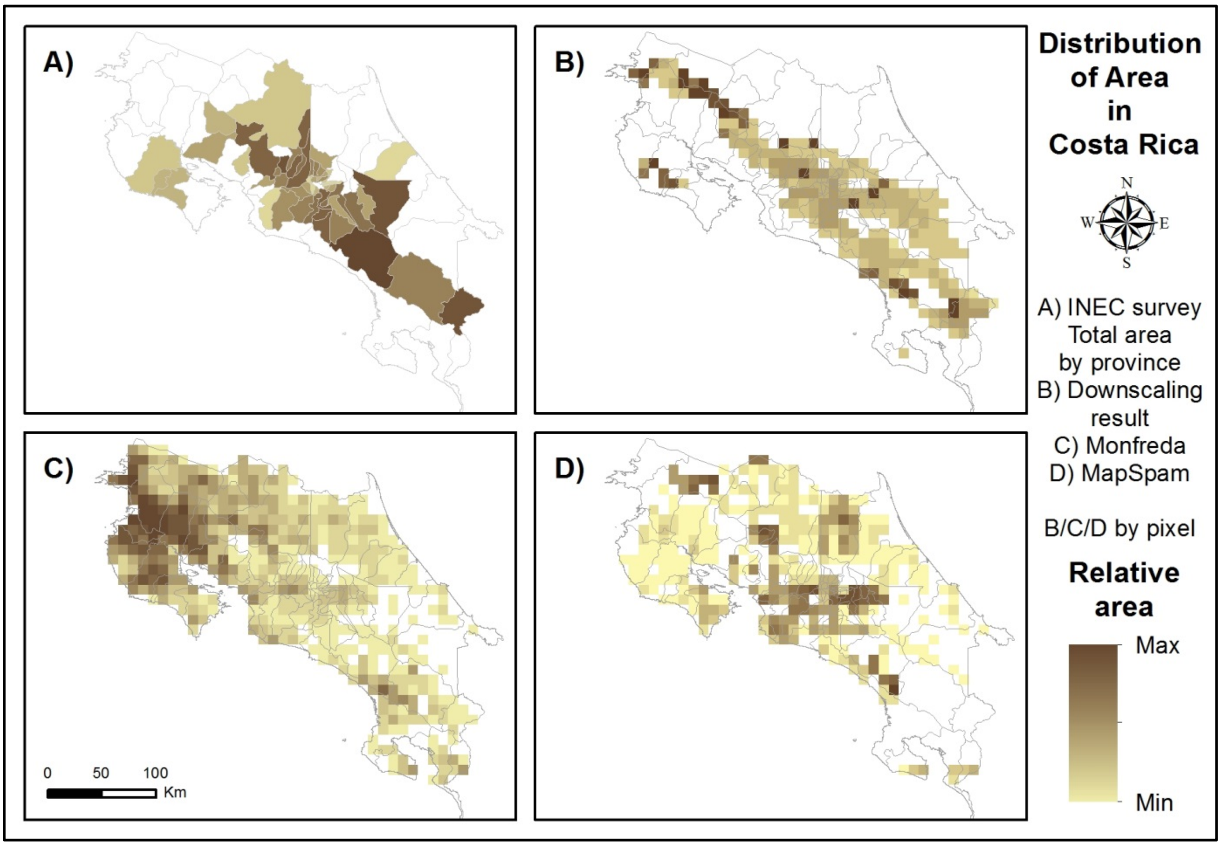
Figure 57. Comparison of area distribution of different data sources (Ethiopia)



Dark brown represents the maximum area, light brown the minimum; A) IFPRI 2006 survey data (adapted from Ruegsegger 2008, original data not available); B) Total area aggregated over Arabica and Robusta; C) Monfreda; D) MapSpam (own representation).

As for Ethiopia results of the different disaggregation approaches differ for Costa Rica. In this region Arabica coffee is produced predominantly under sun and some under biodiverse shade. Subnational data was not included in AgroMaps subnational data. Again, Monfreda placed more coffee area outside of the actual coffee regions according to the survey data, than inside. MapSpam described the major coffee regions somewhat well but omitted area in Southern provinces. My effort described a distribution of coffee area along the central mountain range in the country. Most of the actual area was described well, although area in the North was overestimated (Figure 58).

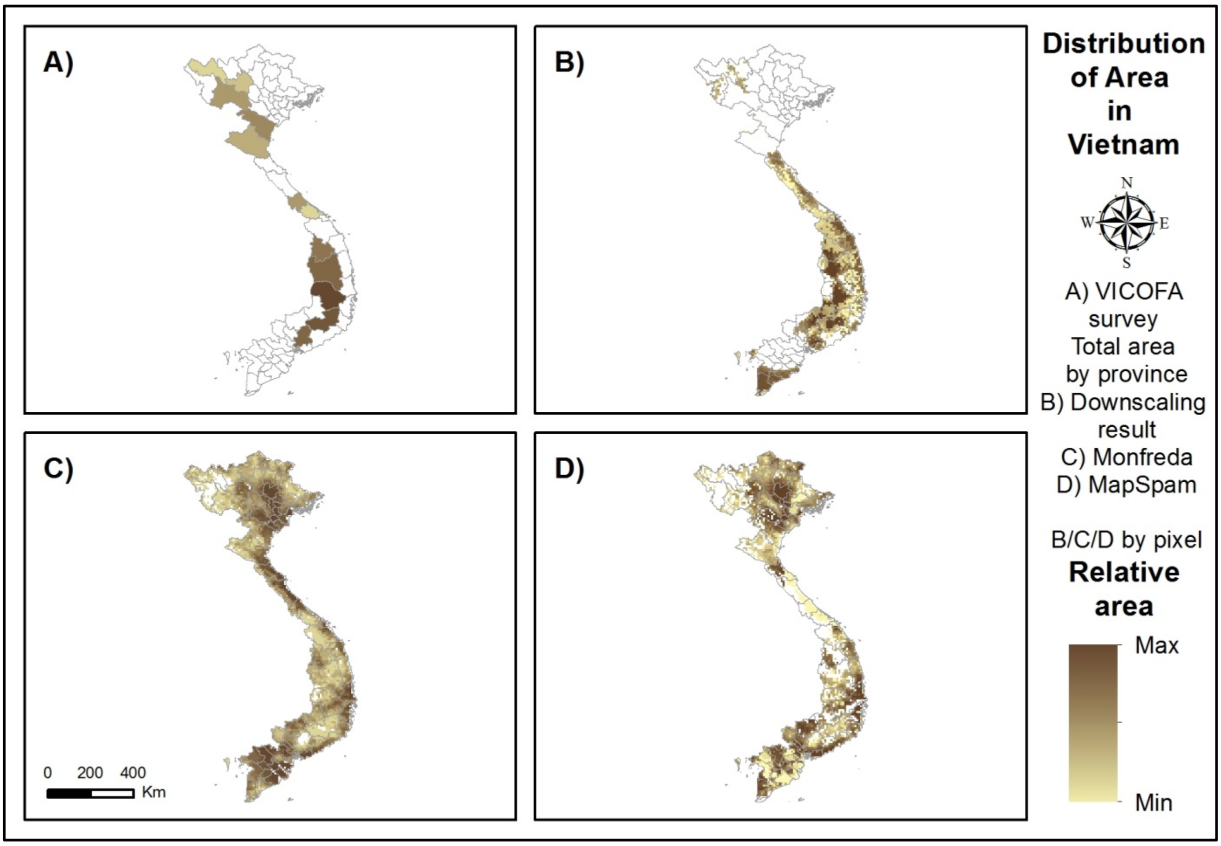
Figure 58. Comparison of area distribution of different data sources (Costa Rica)



Dark brown represents the maximum area, light brown the minimum; A) INEC 2006 survey data; B) Total area aggregated over Arabica and Robusta; C) Monfreda; D) MapSpam (own representation).

Vietnam produces predominantly Robusta under sun and some Arabica. Subnational data was not included in AgroMaps. According to (VICOFA 2004) survey data the largest share of total area can be found within only a few provinces to the South. Without subnational data available Monfreda and MapSpam distributed area for coffee production over the entire country to grid cells with land cover for agricultural uses. My approach reflected the overall pattern of the distribution of coffee within the country somewhat closer. However, substantial area was placed into the southern tip of Vietnam where no coffee appears to be grown (Figure 59).

Figure 59. Comparison of area distribution of different data sources (Vietnam)



Dark brown represents the maximum area, light brown the minimum; A) (VICOFA 2004) survey data; B) Total area aggregated over Arabica and Robusta; C) Monfreda; D) MapSpam (own representation).

Compared to the other efforts the approach presented here thus reflected actual distributions with greater certainty than previous efforts. Especially where no subnational data was available or production is characterized by biodiverse shade this new method is less likely to misplace area to grid cells that are unfeasible. Nevertheless, some area was underestimated, though most error seems to be of the commission type.

4.3 Discussion

Understanding land use change, resource use and the economics of global change processes requires detailed information on the distribution of crop production. The data made available here is a step forward to close the knowledge gap on the physical distribution of coffee production. The methodology derived a prior probability for spatial disaggregation from the crop suitability maps from chapter 3. This innovative approach provided the necessary data to include coffee data in Globiom.

The main result of this section was a database and maps of coffee production area for each coffee producing country by the coffee production systems “Arabica under sun”, “Arabica under biodiverse shade”, “Robusta under sun”, and “Robusta under biodiverse shade”. No reference data for these systems was available so that for comparison with subnational production data the spatially explicit data had to be aggregated by species or a generic “coffee” category. The model results showed reasonable alignment with survey data. Major production regions were well reflected but minor regions were often underestimated. Also, the model allocated area to unfeasible locations in Southern Vietnam.

However, comparing area allocated to grid cells of known coffee occurrences showed that fewer locations were omitted as compared to previous efforts. Nevertheless, also more area was allocated to grid cells of a random background sample for which it is unknown whether these cells are of occurrence or absence type. A visual comparison of this effort with previous disaggregation efforts showed that in countries with a lack of subnational production data in the Agromaps database, this effort is an improvement, while not being worse in countries with detailed subnational data. Beyond the scope of this thesis was a more systematic comparison of data using e.g. a pixel wise Gaussian filtered analysis (Anderson et al. 2015).

This chapter demonstrated the importance of the availability of subnational datasets of crop production. A problem that was encountered was not so much to estimate where coffee is grown, but to know where it is not grown. In general, omission of actual growing areas was quite low, only minor regions were not estimated to hold area. In contrast the climatic suitability model found regions suitable where coffee cannot actually be found according to subnational data. Better climate information, the inclusion of soil data or road access data could allow further improvements of this approach.

Future research should also address the issue of the use of shade as a means to adapt to adverse climate conditions. Here, a uniform distribution of shade and sun systems was assumed across each country. Oftentimes shade is used to adapt production to hot or variable climates. An inclusion of this would further improve the usefulness of this data, especially for climate change adaptation and impact studies. The same holds true for the inclusion of irrigation information. The latter would be the most imminent improvement of this dataset as water use is already a limiting resource in several coffee production regions. Its study is therefore of high interest in the coffee sector.

Despite its shortcomings this disaggregation effort represents a substantial improvement over previous efforts for coffee data even when aggregating into a generic green coffee category. In addition, this is the first globally coherent spatially explicit dataset of a coffee specific disaggregation into production systems. These systems differ fundamentally in their resource requirements, the climatic and socio-economic conditions under which they strive, and also in their typical outputs. These results will be an important step towards an improved spatial analysis of coffee production globally. Furthermore, the methodology presented here has the potential to be applicable to similar crops that are equally underrepresented in subnational datasets such as cocoa, banana or rubber trees.

5 Integrated biophysical-economic assessment of the climate change impacts on global coffee production

In this chapter I will present an integrated climate change impact assessment for the coffee sector using a combination of the previously presented biophysical impact model, the spatially explicit production data and the global recursive dynamic partial equilibrium model Globiom. The Globiom model has been designed to provide policy analysis concerning land use competition between major land-based production sectors such as agriculture, bioenergy and forestry (Havlík et al. 2011). However, like other equilibrium models Globiom previously did not include a model of the coffee sector. This is despite the importance of this crop in regions that are likely to be affected by population growth (Lambin and Meyfroidt 2011), climate change (Vermeulen, Grainger-Jones and Yao 2014), and increasing livestock product demand (Phalan et al. 2013). These factors are generally seen to affect food security, and to drive direct and indirect land use change (Phalan et al. 2013). The explicit inclusion of coffee production into Globiom is therefore likely to give additional insights that neither the Globiom model without coffee, nor the coffee biophysical impact model from chapter 3 could provide in isolation.

This chapter is aimed at examining the effects of climate change when accounting for demand side effects by means of an economic model. Foremost, Globiom will be used to answer the same questions as in the chapter on the biophysical impacts. Previous research (chapter 2.2), and the research presented in chapter 3, suggested that climate change would result in latitudinal migration (Zullo et al. 2011), altitudinal migration (Simonett 1988; Schroth et al. 2009), or regional migration (Schroth et al. 2014) of coffee production. Especially the question whether Robusta will be able to substitute Arabica coffee could not be answered in chapter 3: While important Arabica producing regions like in Brazil will be negatively impacted by climate change (chapter 0), the most significant negative impacts on suitability for Robusta production were projected in the Congo basin (chapter 0). This region however does not produce large quantities of coffee (chapter 4.2.3). Together with the generation of

Authorship:

The Globiom model, its data and the approach to climate impact modeling were designed, written, and published by the IIASA colleagues. Inclusion of coffee specific data and scenarios was my own work.

spatially disaggregated area data in chapter 4 with the inclusion of the Globiom model it is possible to examine feedbacks between the impacts on the two production systems. This analysis will then be supplemented with economic indicators that can only be assessed using an economic model, i.e. demand growth, changes in relative prices and consumption patterns.

This chapter furthermore seeks to examine the effects of the inclusion of climate change impacts on coffee production in Globiom. While coffee production takes place on relatively small areas compared to staple crops, at its origins it occupies relevant area. Therefore, a question will be whether at its origins these impacts on coffee also have indirect effects on the production of other crops.

Last, the usefulness of the Globiom model with coffee is demonstrated by modeling a scenario in which a potential relative winner of the climate change impacts on coffee seeks to benefit from opportunities. East Africa will be affected less by climate change than other regions (chapter 3). However, productivity in Africa is generally low (chapter 2.3.3). Here, a simple scenario was used in which Ethiopia closes its current yield gap in Arabica production to Brazil by 2050. It is then discussed whether the net present value of the investment changes as a result of climate change.

The following will initially describe the integrated modelling framework Globiom with a focus on the implementation of the climate change scenario. Then the necessary steps to include coffee production in Globiom are presented before a brief outline of the method used to derive results. These results of the integrated modeling framework will also be compared with the results derived from the analysis of biophysical climate change impact model. However, results from chapter 3 were not directly comparable. For Globiom only data for the climate change impacts on crops in the A2 scenario of the 4th IPCC report was available (Mosnier et al. 2014). For consistency previous analysis therefore had to be partially repeated for this scenario.

5.1 Integrated modeling framework

For this chapter I followed the integrated modeling concept as outlined in Figure 1. Biophysical impact models driven by global climate models were used to change the equilibrium in the global recursive dynamic partial equilibrium model Globiom. A general description of Globiom can be found in (Havlík et al. 2011). A description of the effect of demand side changes can be found in (Schneider et al. 2011) where the model is used to research the effects of population growth on land use change.(Mosnier et al. 2014) used it to estimate the cost of a consumer support policy to mitigate climate change impacts on food security.

Globiom uses the MapSpam database of crop distribution (You et al. 2012) for baseline data input, and the EPIC model (Williams and Singh 1995) to estimate yield potentials. Future productivity changes were estimated based on ordinary least squares regression on FAO time series for the years 1980 to 2010. Demand was modeled based on data from (Alexandratos et al. 2006). To include a coffee model into Globiom it was therefore necessary to model the current distribution of coffee production, estimate current and future yield potential and future productivity and demand changes.

The yield potential changes caused by climate change were modelled for the CNRM CM3, MRI CGCM 2.3.2 and UKMO HadGem1 global climate models for the A2 scenario (often referred to as “business as usual scenario”) by EPIC in the study by (Mosnier et al. 2014). Because this data was used to include climate change impacts on the crops already implemented in Globiom, these GCMs were also used for the coffee model in this chapter.

5.1.1 Integrated climate change impact modeling with Globiom

For climate change impact modeling with Globiom the entire chain from SRES-GCM-biophysical model-economic model had to be integrated. Globiom is recursive dynamic. It is calibrated with baseline year data and iteratively solved for each scenario period. Data therefore had to be provided in ten year steps from 2000 to 2050. Some data is continuous between these time steps, e.g. population trajectories. However, yield potential data for each period was based on mean GCM outputs for (2040 to 2069 = “2050s”) and interpolated between periods.

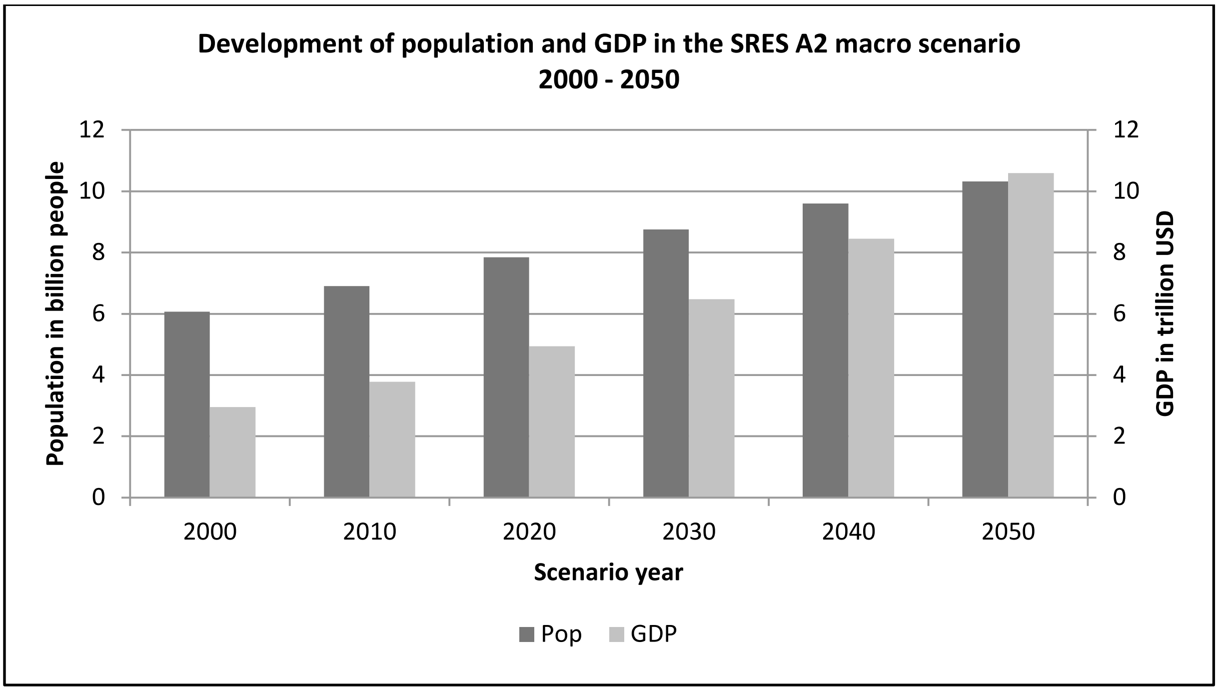
The assumptions on population and GDP growth from the SRES drove demand side changes in Globiom. On the supply side SRES fed into the global climate models and the outputs of GCMs changed the solution of the EPIC biophysical yield potential model. Changes in yield potential in turn provided the change on the supply side of the Globiom model. In the following the assumptions made at each step are described.

5.1.1.1 Macro trends of population and GDP

Development trajectories for population and GDP development were taken from the A2 SRES scenario. This scenario represents emissions of "a very heterogeneous world. The underlying theme is self-reliance and preservation of local identities. Fertility patterns across regions converge very slowly, which results in high population growth. Economic development is primarily regionally oriented and per capita economic growth and technological change are more fragmented and slower than in other storylines" (Nakicenovic and Swart 2000, Box TS-1).

This macro scenario projects a population increase to 8.6 billion people by 2030 and 10.3 billion in 2050. India overtakes China as the most populated country in the 2020ies and Western African population is projected to triple from 300mio to 900mio by 2050. At the same time global GDP grows from 35 trillion USD to 123 trillion USD. The most drastic economic growth is projected for China which would reach 20 trillion USD by 2050, nearly matching the USA as the two largest economies. Global GDP per person thus doubles from year 2000 USD5683 to USD11,974 in 2050 (Figure 60).

Figure 60. Development of population and GDP in the SRES A2 macro scenario (2000-2050)



Dark grey bars shows the projected population increase between 2000 and 2050; light grey bars the GDP development (own representation of Nakicenovic and Swart 2000).

The income effect of on future demand followed FAO assumptions as in (Alexandratos et al. 2006). Demand is affected in this scenario through changes in population size and through region specific income elasticities of demand. Because global GDP was projected to increase in most regions the total calorie demand will rise to somewhere between 3000 and 4000kcal/capita/day. The composition of calorie demand differentiated calories from animal sources and crop sources. From one period to another in Globiom shifters were applied for calorie consumption to include changes in dietary patterns. On global average the ratio of animal calories to vegetal calories was projected to increase from 1:4 to 1:3. Shifters were region specific so that animal based calorie consumption increased much more in regions like China than in Western Europe (Schneider et al. 2011).

5.1.1.2 Technical progress scenario

Technical progress in Globiom was modeled for each of the 30 regions based on FAO yield data for the period 1980 to 2006. A regression model was used to extrapolate the development of yields during this period into future periods. The resulting scenario of technical progress corresponded to an average annual growth rate of approximately 1% for developed countries, and 1.3% for developing countries (Havlik et al. 2012). The yield change was applied as a shift factor for each period relative to the baseline yield modeled by EPIC.

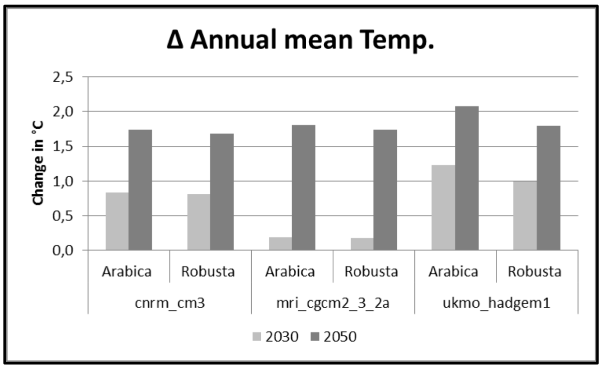
5.1.1.3 Scenarios for climate change

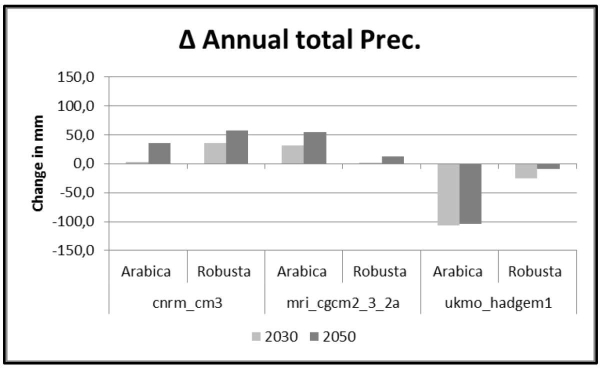
In line with the macro scenario the global climate change models used here were also driven by the SRES A2 scenario. The scenario corresponds to high temperature increases of about 2.0°C until 2050 and 3.4°C until 2100 (Solomon et al. 2007) when regarding the multi-model mean of 17 GCMs included in the AR4 report. However, disagreement between GCMs is considerable regarding changes in precipitation. Therefore a set of GCMs was chosen that is representative of the range of projected changes in the Climate Moisture Index (CMI), an indicator of aridity. Taking into consideration a global wet, global dry and a global mid-range climate by 2050 resulted in three GCMs: MRI-CGCM2.3.2 (“MRI”) as a wet scenario, UKMO-HadGEM1 (“UKMO”) as a dry scenario, and CNRM-CM3 (“CNRM”) as the mid-range scenario (Mosnier et al. 2014).

To support the interpretation of impact projections the climatic changes at coffee occurrence locations (chapter 3.1.1) are briefly presented here because regional climatic changes often differ from global trends. The changes at coffee occurrence locations in the location database could therefore be different than the global trends.

Until 2050 the increases in annual mean temperature will be somewhat similar across all three GCMs at occurrence locations. All GCMs projected an increase of about 1.7°C at coffee locations. The only exception was the UKMO GCM at Arabica locations where it projected an increase of 2.1°C. Until 2030 the disagreement was higher. MRI projected a mere 0.2°C increase, UKMO ~1.1°C at coffee locations (Figure 61).

Figure 61. Climatic change of annual means at coffee occurrence locations 2030/2050





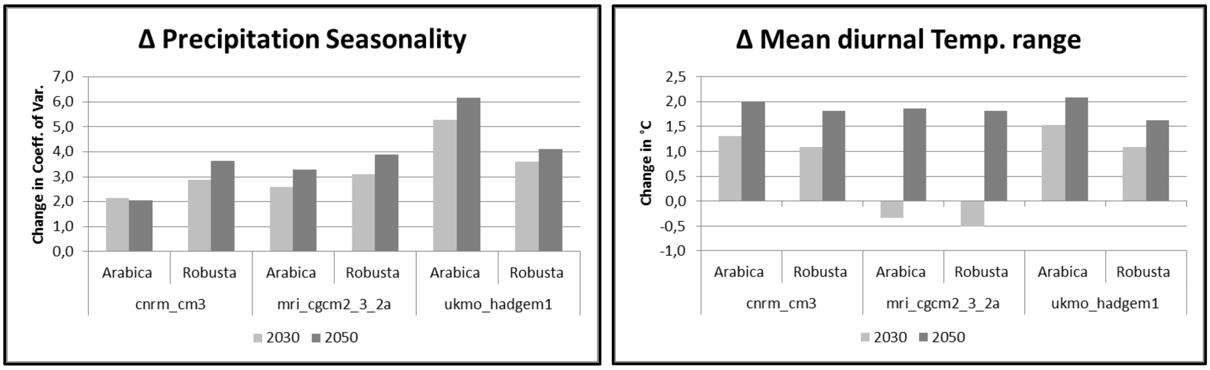
Annual mean temperature and annual total precipitation at coffee occurrence locations compared to historic climate; light grey - 2030; dark grey – 2050 (own data).

Projected changes of annual total precipitation did not differ much between CNRM (global mid-range) and MRI (global wet) for both time slices 2030 and 2050, in all cases the increases ranged between 2 and 57mm of mean annual total precipitation. The UKMO GCM (global dry) however suggested decreases of up to 107mm at Arabica locations and 25mm at Robusta locations (Figure 61).

These indicators describe general changes at coffee regions. In chapter 3.2.4 we analyzed the variable importance for the coffee suitability models. For Arabica the precipitation seasonality (the coefficient of variance of monthly precipitation) and for Robusta the mean diurnal temperature range (mean monthly max. temp – mean monthly min temp.) were found to be most influential (Table 12).

The precipitation seasonality will change according to all GCMs. The most drastic changes were projected to happen until the 2030ies with little additional change until the 2050ies. The highest change was projected at Arabica locations in the UKMO model (Figure 62). The mean diurnal temperature range at coffee locations will increase by approximately 2°C until 2050 in all GCMs with 1°C increases by 2030. An exception is the MRI scenario that projected modest decreases in the diurnal temperature range by 2030.

Figure 62. Climatic change variability indicators at coffee occurrence locations (2030/2050)



Changes in precipitation seasonality and mean diurnal temperature range at coffee occurrence locations compared to historical climate; light grey - 2030; dark grey – 2050 (own data).

5.1.1.4 Biophysical impacts of climate change on other crops

In Globiom yield potential data was modeled by EPIC (Williams and Singh 1995) for each crop and related production systems. The data is entered in Globiom at SimU level. To simulate climate change impacts the EPIC model was run for both current and future conditions. Differences in yield potential were applied as relative yield shift factors for each SimU during Globiom scenario simulations. EPIC used the temperature, precipitation and CO_2 levels as provided by the GCMs as inputs. The model accounted for the anticipated fertilization effect of elevated CO_2 in the atmosphere. EPIC's outputs were potential evapotranspiration, crop yield potential, and input requirements for each crop, production system and climate model on global scale.

Using a similar model set up and scenario (Mosnier et al. 2014) investigated the consumer cost of climate change adaptation. They reported that on global level the EPIC biophysical model projected reductions in total calorie production by -2% or increases by up to +0.03%. Impacts were shown to be heterogeneous across regions. Independent of the GCM scenarios yield potentials were shown to decrease by 2050 in Southern U.S.A., the sub-Saharan Sahel zone and Northern Chinese plains. The UKMO and CNRM scenarios also projected losses in yield potential for India, while the "wet" MRI scenario projected increases in yield potential in this region. Across GCMs Europe, Southern Brazil and Argentina and southern Africa were projected to experience yield increases (Mosnier et al. 2014) Figure 2, p. 35).

5.1.2 Inclusion of coffee in Globiom

Inclusion of coffee in Globiom required the development of a demand scenario and the provision of data on SimU level for yield potential and harvested area.

5.1.2.1 Coffee demand scenario

The demand scenario for coffee assumed a concave GDP per capita elasticity of demand function. Elasticities were low for both low income countries and high income countries. The highest income elasticity of demand was assumed for economies that are shifting from middle income to high income as a result of the macro scenario. Such a demand scenario cannot be supported empirically for all countries (Webb and Hall 2009), but nevertheless appeared feasible given that high income markets were reported to be stagnant, while emerging markets are growing rapidly (Lewin et al. 2004).

The empirical estimation of elasticities of demand for coffee is confounded by the price regulation regime of the ICO that lasted until 1989, the subsequent period of historically low prices for coffee, and a lack of consistent data. There is some evidence that income increases in high income countries will not increase demand. But, increased demand can be expected from income increases in other countries (Webb and Hall 2009). This is supported by recent rapid increases of coffee consumption in emerging economies in Asia, Eastern Europe and the former Soviet Union (Lewin et al. 2004).

In Globiom, the definition of income categories followed the WorldBank definition from 2004 (Soubbotina 2004). Low income countries had a GDP per capita of less than \$755, lower middle income countries less than \$2995, higher middle income countries less than \$9,265, above which countries were considered high income (Soubbotina 2004). For the coffee demand scenario low income and lower middle income countries were placed into the same category, where growth in GDP per capita has little demand effect. China and India were in this category, but also the greater part of Africa and Asia. Equally low market expansion was assumed for countries in the mature market category with GDPs per capita greater than \$20,000. This threshold approximates the one found by (Webb and Hall 2009), and placed Northern and Western European countries, the U.S.A. and Japan in this category, where demand growth stems almost entirely from population changes. For middle income countries below \$6,265 and high income countries above \$12,265 GDP per capita an intermediate elasticity of demand of 0.8 was assumed. In the base year 2000 mostly South American countries can be found in this category. In the same year Mexico and South Korea were the only regions in the emerging market category with a high income elasticity of demand of 1.5 (Table 18). As the macro GDP growth scenario progressed other regions entered this category, e.g. China, Russia, and the remainder of Europe.

Table 18. Coffee demand scenario assumptions for GDP per capita income classes and assumed income elasticities of demand.

Income class	Definition in USD	Income elasticity
Low income	<2,995	0.1
High middle income	<6,265	0.8
Emerging markets	<12,265	1.5
High income	<20,000	0.8
Mature market	>20,000	0.1

(Own data)

As for the other crops, price elasticities of demand for the 30 Globiom regions were obtained from (Seale, Regmi and Bernstein 2003) (Table 19). Globiom used a maximum price elasticity of -0.3, so that values above this were reduced. In high income countries the price elasticities of demand are generally low, so that for these regions the values were assumed to be -0.3. Especially very low income countries in Africa had extreme elasticities beyond -1.

Table 19. Globiom regions and price elasticities of demand for coffee

Region	Price elasticity of demand	Region	Price elasticity of demand
ANZ	-0,33	**RCAM**	-0,73
BrazilReg	-0,71	**RCEU**	-0,75
CanadaReg	-0,3	**ROWE**	-0,3
ChinaReg	-0,9	**RSAM**	-0,7
CongoBasin	-1	**RSAS**	-0,89
EU_Baltic	-0,68	**RSEA_OPA**	-0,73
EU_CentralEast	-0,63	**RSEA_PAC**	-0,9
EU_MidWest	-0,34	**SouthAfrReg**	-1,24
EU_North	-0,36	**SouthKorea**	-0,47
EU_South	-0,41	**EasternAf**	-1,34
Former_USSR	-0,78	**SouthernAf**	-1,1
IndiaReg	-0,89	**WesternAf**	-1,27
JapanReg	-0,31	**TurkeyReg**	-0,67
MexicoReg	-0,65	**USAReg**	-0,11
MidEastNorthAfr	-0,76	**RCAM**	-0,73
Pacific_Islands	-0,67	**RCEU**	-0,75

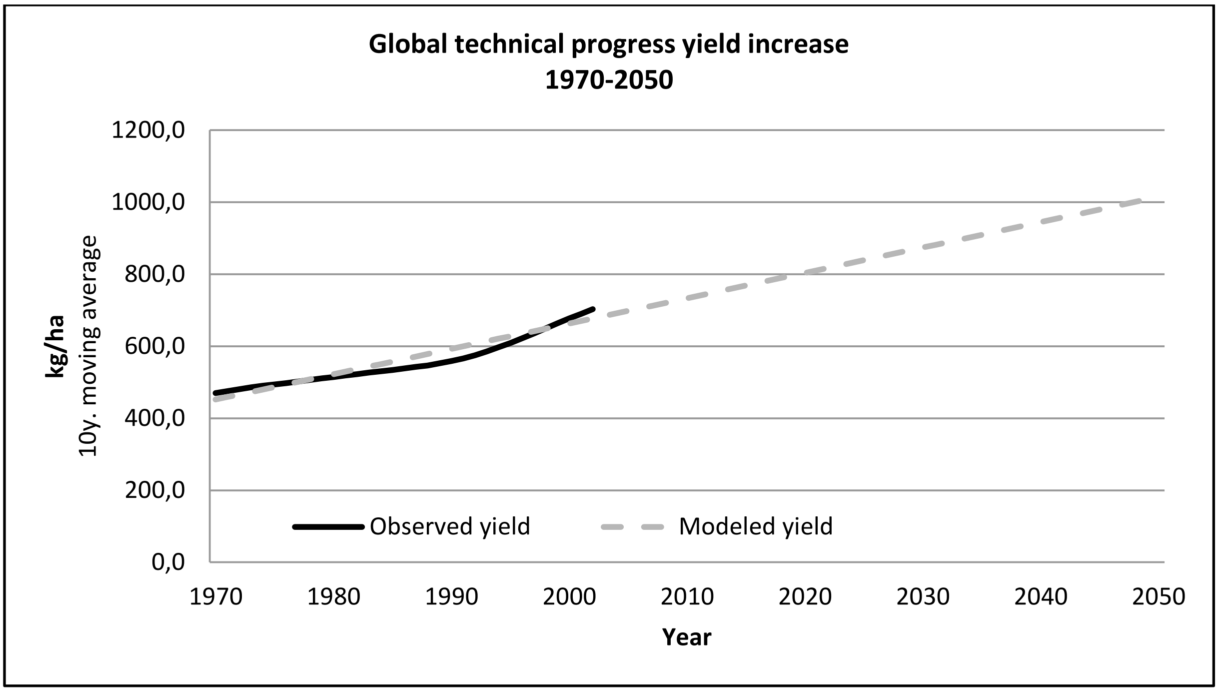
(compiled from Seale et al. 2003)

5.1.2.2 Coffee technical progress

In Globiom technical progress was modeled as a linear extrapolation of historical yield data from FAO for each region. Therefore, technical progress for coffee production was modeled in a similar fashion but as a global trend.

FAO yield data was aggregated in the "green coffee" category which comprises both species. This data was incomplete and calculated based on reported, interpolated, or estimated harvested area and total yields. Since the liberalization of the coffee market in the 90ies yields have increased substantially (Figure 63). This was likely driven by increased competition. Nevertheless, not all regions could benefit, e.g. African countries have seen stable or declining yields over the past five decades (Figure 30). Basing an extrapolation of technical progress on the 1980-2006 period as was done for the other crops in Globiom would thus overestimate future yield increases for competitive regions, and underestimate future yield increases in Africa. Therefore, here a globally homogenous yield increase for both species was assumed based on a linear extrapolation of FAO yield data from the 1965-2006 period (Figure 63).

Figure 63. Development of global historic coffee yields (1970-2005) and technical progress scenario until 2050



Extrapolation of historic yield increases (solid black line) (FAO 2014b) as a technical progress scenario (dashed grey line) for both coffee species (own representation).

5.1.2.3 Baseline area data

The baseline area data for model calibration was produced as described in chapter 4.1.3 with minor adaptations. It was not differentiated between production under biodiverse shade and under sun. Also, the disaggregation was not conducted to grid cells but to SimUs.

In chapter 4 machine learning methods from chapter 3 were applied to distinguish climate in coffee production countries from the climate at geo-referenced locations of coffee production to derive spatially explicit probabilities of presence of coffee production sub-nationally. Using a cross entropy approach harvested area statistics for *C. arabica* and *C. canephora* were allocated to areas with suitable climate and probable land use characteristics. This was subject to the condition that allocated area could not exceed maximum concentrations observed in important coffee regions. The result was a spatially explicit database of coffee production area for each coffee producing country by the coffee production systems “Arabica under sun”, “Arabica under biodiverse shade”, “Robusta under sun”, and “Robusta under biodiverse shade”.

To include this data in the Globiom model structure the harvested area data was disaggregated into SimUs rather than grid cells. To do so, prior probabilities were averaged for each SimU before disaggregation. Furthermore, the disaggregation model was amended with the condition that for each downscaling cell area could not exceed area available in the Globiom database. Area under shade and area under sun were then summed for each species. The result was a database of the baseline distribution of coffee production for the two species that matched the database used by Globiom without causing data conflicts.

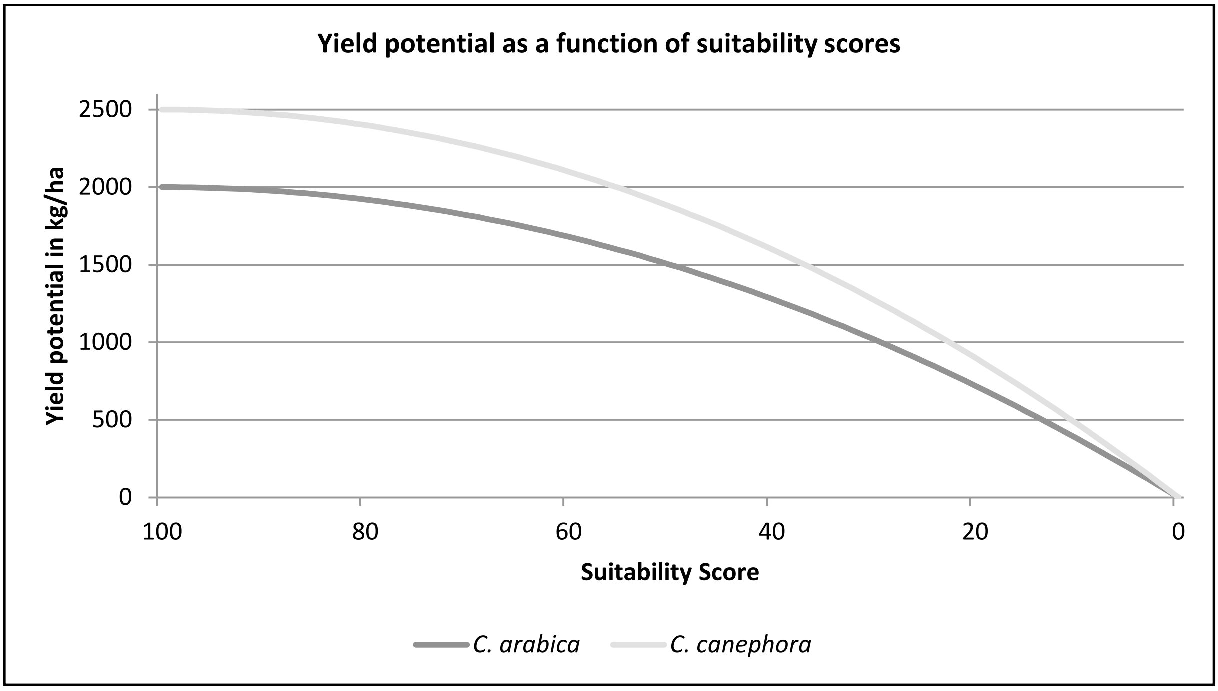
5.1.2.4 Climate change impacts on coffee yield potential

Analogous to the EPIC yield potential data for other crops yield potential data for all time slices for the two coffee systems was required. Therefore the suitability scores from chapter three were translated into estimates of yield potential. First suitability maps were modeled, then a function to estimate the yield potential from suitability was applied and last, the impacts of climate change on coffee yield potential were compared with the EPIC results for other crops.

Climate change impacts were modeled as described in chapter 3 but using the AR4 GCM data (CNRM, MRI, UKMO) for 2030 and 2050 from the A2 scenario. In chapter three machine learning classification algorithms were trained on a global coffee occurrence location database using parameter spaces. Parameters were chosen within reasonable ranges to avoid both overtraining (excess of specificity) and lack of specificity in overgeneralizing models. The resulting 135 models were extrapolated to raster data for current conditions from the Worldclim database and downscaled AR5 GCM outputs. The multi model mean results showed a drastic reduction in global suitability for both *C. arabica* and *C. canephora* across emission scenarios by 2050.

Based on suitability surfaces yield potential for current and future conditions was estimated. To prepare climate change impact data for this chapter the classification models from chapter 3 were applied to downscaled 5'arcmin GCM outputs for the A2 scenario by 2030 and 2050. This yielded suitability surfaces for future climate conditions for the GCM's the CNRM CM3, MRI CGCM 2.3.2 and UKMO HadGem1. Globiom uses EPIC crop model outputs on a 30'x 30' grid to model current and future yield potential. Here, from global suitability maps an estimate of yield potential was derived. In the literature stated values for maximum yields under optimal management for Arabica were estimated to be around 2000kg/ha and about 2250kg/ha for Robusta systems (Wintgens 2009). Here, it was assumed that such yields can be achieved at locations with optimal climatic suitability. At locations with marginal climatic conditions the yield potential was assumed to be approximately half of the optimal value (Figure 64).

Figure 64. Yield potential modeled as a function of suitability



Dark grey line *C. arabica* and light grey line *C. canephora* (own data and representation).

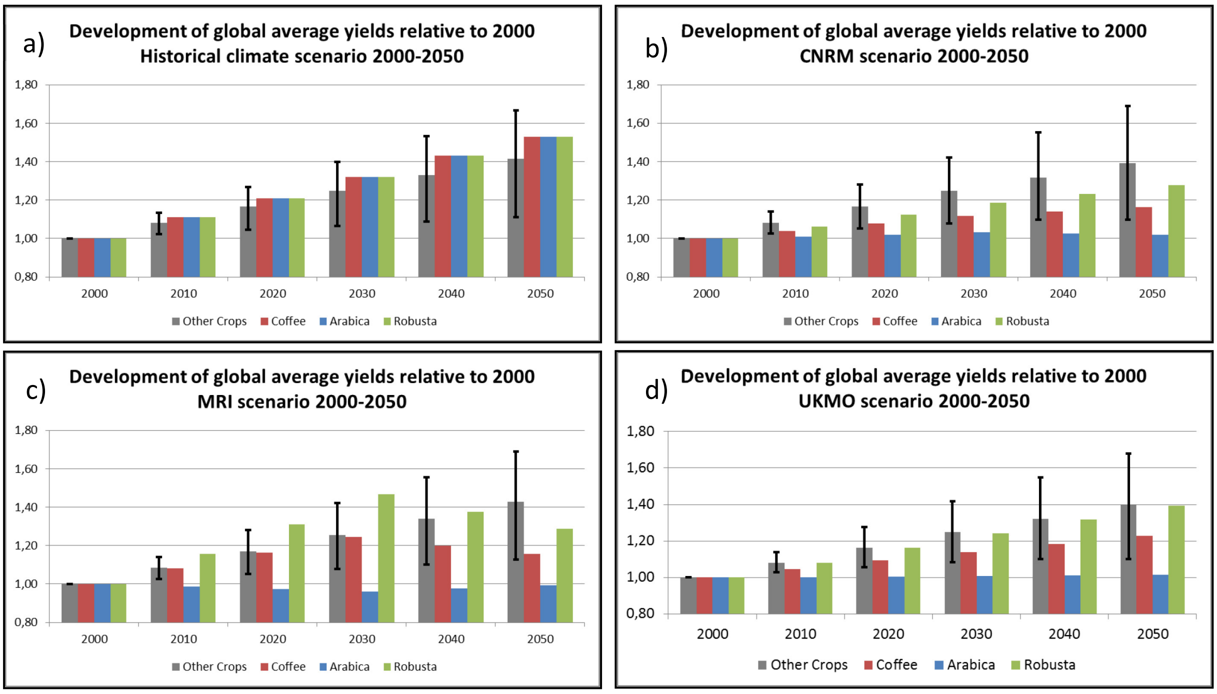
The classification model mean output based on bioclimatic variables was an estimate of the likelihood that the climate at a specific location is suitable for a species. I used equation (8) to estimate spatially explicit yield potentials for the Arabica and Robusta systems both for current and future climatic conditions based on the suitability score.

$$Y_{ij} = Y_{max,j} \times \left(1 - \left(1 - Suit_{ij}\right)^2\right) \qquad \text{(Eq. 8)}$$

Y_{ij} is the yield potential in cell i for system j, $Y_{max,j}$ the maximum achievable yield for system j and $Suit_{ij}$ the suitability index from the multi model mean in cell i for system j. The resulting function is plotted in Figure 64. This function was applied to current and future (2030/2050) suitability surfaces. The respective data for the scenario time slices 2010, 2020 and 2040 was produced by interpolation of the impact values of neighbouring time slices. These changes in yield potential were applied to the baseline yield potential as relative impact shift factors to each SimU during each model simulation repeat.

In the no climate change reference scenario biophysical yield potentials increased over time as an extrapolation of historical yield developments. Yields for both coffee production systems, Arabica and Robusta, increased at same rate by 53% until 2050. Yields of other crops were projected to increase by 11% to 67%. The yield increase of coffee systems was thus higher than the average yield increase of 41% (Figure 65). For the other crops in Globiom climate change impacts were modeled by EPIC. The average relative impact on yield potential was -1% (CNRM, UKMO) or +1% (MRI) compared to the reference scenario. For some crops yield potentials rose by up to 10% (dry beans) or impacts were higher (-10%; Sugarcane) (Figure 65b-d). This is in line with the results from a similar model set up described in (Mosnier et al. 2014). There, instead of yield potential calorie production was considered and reported biophysical impacts vary between +2% and -3%, depending on the GCM.

Figure 65. Development of global average yields relative to baseline yields until 2050



a) No climate change scenario and b-d) with climate change; Grey bars represents the average of all other crops, the error bars the range; red bars are the average for coffee, blue are Arabica only; and green Robusta only (own data and representation).

Negative impacts by climate change on the average yield potential of the two coffee systems were higher than for other crops. By 2050 the average yield potential was approximately 20% lower than in the NoCC reference scenario in all GCM scenarios. Arabica was projected to be impacted by yield potential reductions of ~34% compared to the reference scenario. Impacts on Robusta were less drastic and varied between -9% and -17% depending on the GCM. In the MRI scenario the Robusta yield potentials could be positively impacted by +11% until 2030, but would then again be reduced by 16% until 2050. The high impacts on Arabica production resulted in a net zero yield improvement compared to the base year 2000, i.e. yields would not be improved despite technical progress (Figure 65b-d).

5.1.3 Estimating the impacts of climate change on global coffee production

To estimate the impacts of climate change on global coffee production beyond the biophysical impacts on yield potential the spatial distribution of coffee production by area and produced quantity was considered. To analyze the impacts on other crops production and price indices were used in two parallel model runs: once with shifts in coffee demand and supply, once without. Finally, an example application is presented.

5.1.3.1 Spatial distribution of coffee production

Globiom is spatially explicit. The combination of the resulting area distribution data as modeled by Globiom with the future yields allowed the elaboration of maps of the distribution of coffee future production. By comparison of the GCM scenarios with the scenario with historical climate the relative climate change impacts could be analyzed. Regions that will lose production because of climate change, and regions that gain production were identified this way.

Additionally, the result of the Globiom area optimization was compared with the outcome of the biophysical impacts as analyzed in chapter 3.1.7. Impacts were compared across latitude and altitude classes. Suitability scores and the area solution acreage were summed up across 1° latitude classes and 100m altitude classes for current climate conditions and for the climate change scenario solutions.

5.1.3.2 Total production and price index

As a measure of impacts on other crops the total global crop production quantity and a price index was used. For each GCM the total supplied quantity from crop production activities was summed up over all regions for crop production. The price index for each scenario was based on the weighted mean price across all regions and crops. For both indices coffee was excluded and regarded separately.

5.1.3.3 Scenario without coffee

While acreage used for coffee production is small in comparison to staple crops, in certain regions it occupies relevant acreage. Therefore the inclusion of coffee in Globiom could alter the implications for food security as modeled by Globiom.

To compare the impacts of the inclusion of coffee in the model Globiom was calibrated and solved as before but without changes in coffee demand and supply. Demand remained constant at base year levels for each region and did not increase with increasing population. Neither technical progress nor climate change impacts factor were applied so that yields remained constant at base year levels.

As before, total production and prices were used as a measure for impacts on other crops. For important coffee production regions differences between the supplied quantity and price index with or without coffee were described.

5.1.3.4 Opportunity scenario for Ethiopia

The distribution model in chapter 3 showed that Eastern Africa could be a potential opportunity site for Arabica coffee. The suitability model demonstrated that in this region suitability losses will be less negative than in other major Arabica growing regions. However, the comparison of the results of chapter 3 with the findings of (Davis et al. 2012) suggested that the climate in the region will change its fundamental characteristics for coffee production. The implication was that the production practices in Eastern Africa will have to be adapted to the novel conditions to seize the opportunity.

Here, the effects of a policy for Ethiopia were investigated that results in an additional technical progress of 1% per year and costs 15mio USD per year, e.g. a coffee growers federation that funds coffee research. Such an increase in yields would close the baseline year 2000 yield gap between currently unproductive Ethiopia and highly productive Brazil by 2050.

To implement the policy in each scenario year an additional shift factor was applied to the yield potential in the Ethiopia region, after application of the technical progress and climate change impact yield shifters. Evaluation of the policy was based on the net present value of the additional income generated in the 2000-2050 period by the policy compared to the GCM and NoCC scenarios without the policy. Following a study on the value of a crop breeding program in Ethiopia (Hein and Gatzweiler 2006) a discount rate of 10%/year was assumed.

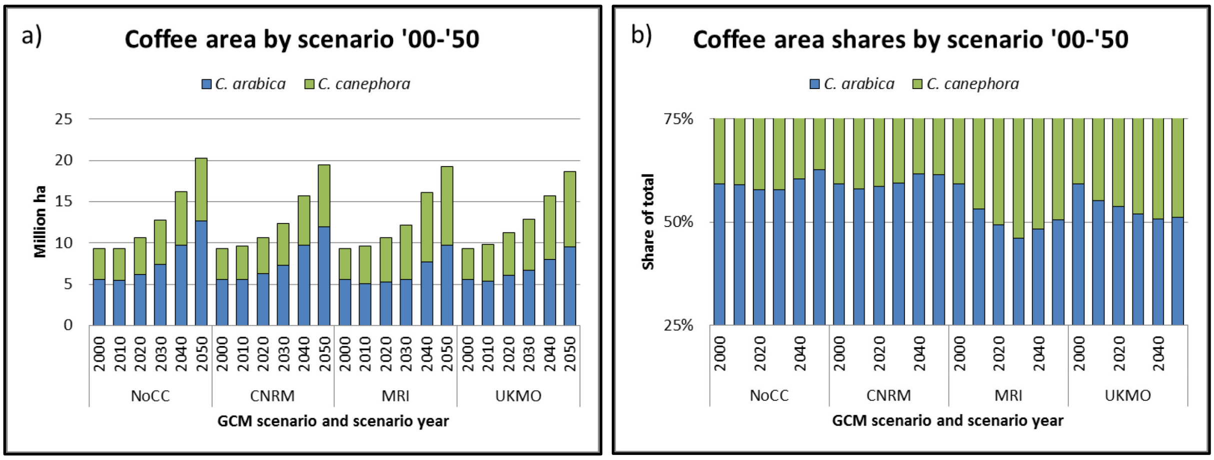
5.2 Results – The impacts of changing paradigms on future coffee production

First, gross impacts of climate change on coffee production will be presented. The differences between the distribution of suitability and harvested area will be compared to explore the impacts of climate change on latitudinal migration, altitudinal migration and regional migration. Mapping the spatial distribution of future coffee production gave further insights. Finally, the effects of including coffee in Globiom on other crops were evaluated and the results of the opportunity scenario for Ethiopia are presented.

5.2.1 Total production and system prevalence

Total area increased in all scenarios from about 10mio ha worldwide to approximately 20mio ha from the base year 2000 to 2050 (Figure 66a). Differences could be observed in the production mix between Arabica and Robusta. In the scenario without climate change impacts (NoCC), the share of Arabica area increased from currently 59% to 63%. With climate change the relative share of Arabica area was projected to decrease in the MRI and UKMO scenarios from 59% to 51%. In the CNRM scenario the relative share increased to 61% (Figure 66b).

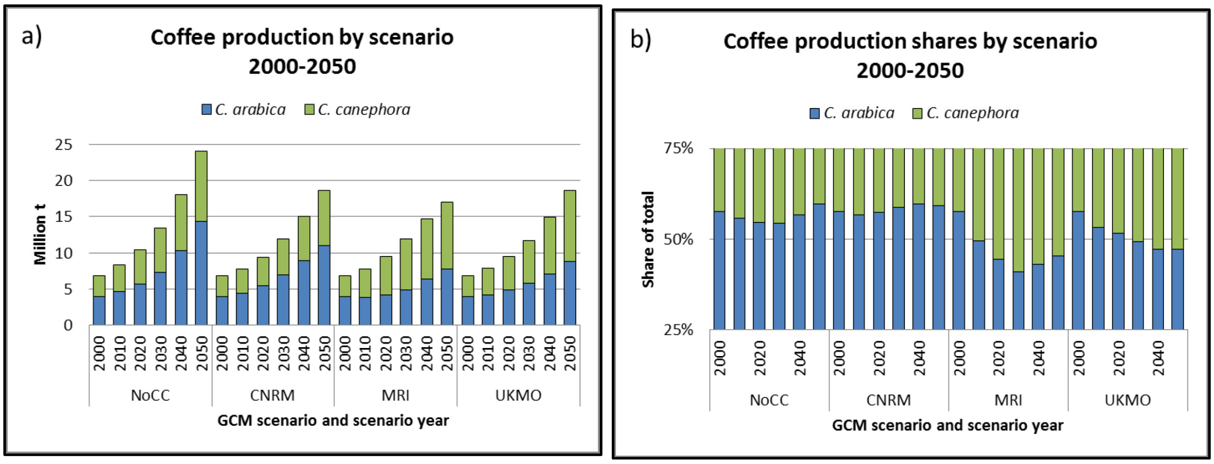
Figure 66. Total area for coffee production by scenario and shares of total area (2000-2050)



a) Total area for coffee production by scenario, split into area planted with Arabica and Robusta in million ha; and b) species share of total area in percent by scenario (own data and representation).

While total area increases were roughly the same in both the reference scenario and the climate change scenarios, this was not the case for total coffee production. In the reference scenario total coffee production was projected to increase 3.5-fold to nearly 25 million tons by 2050. In the scenarios with climate change this increase was reduced to a 2.5-fold increase to approximately 19million tons (Figure 67a). This corresponded to a relative loss of production of 22%-29% in the GCM scenarios compared to the reference scenario without climate change.

Figure 67. Total coffee production by scenario and shares of total production (2000-2050)



a) Total coffee production by scenario, split into Arabica and Robusta in million t. and b) species share of total production in percent by scenario (own data and representation).

In the NoCC reference scenario the relative share of Arabica based production increased less than the relative share of area: from 58% to 60% instead of 63% (Figure 67b). As for the area, in the CNRM scenario ("mid-range") the share of Arabica production remained roughly the same. In the UKMO and MRI scenarios, however, the decrease of the Arabica production was more pronounced than when looking at area only. The lowest relative Arabica production was projected in the "wet" MRI scenario by 2030 with only 41%.

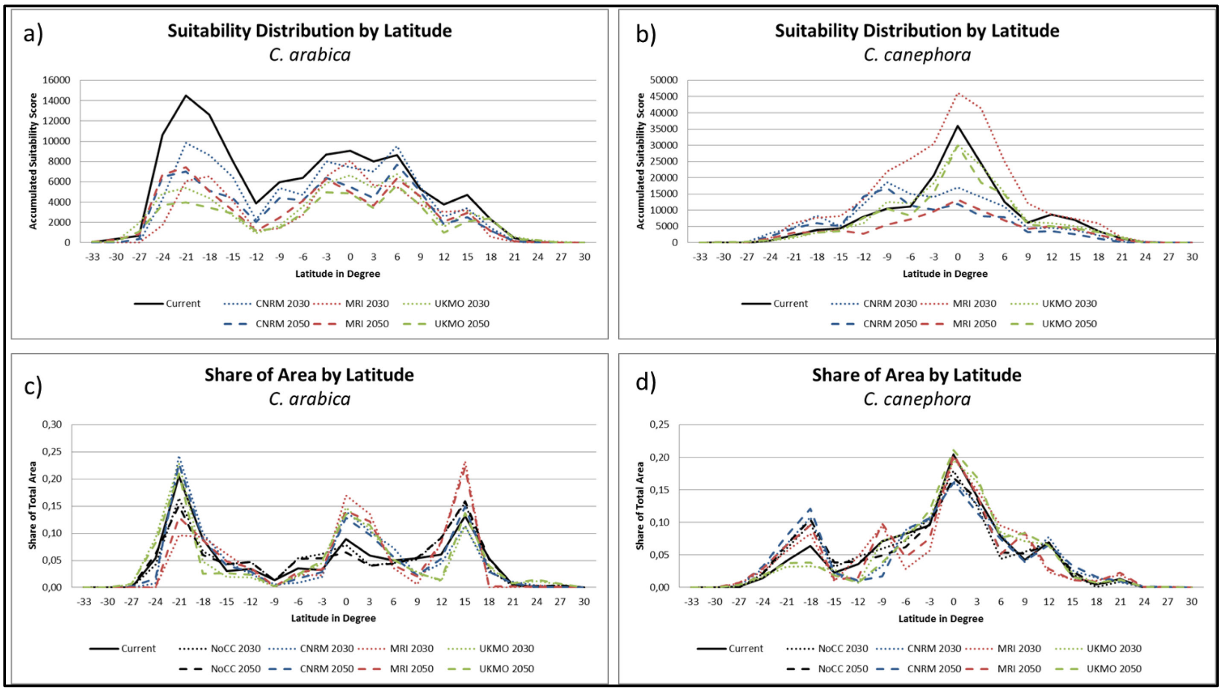
5.2.2 Distribution of suitable and harvested area

A key question raised in previous studies was how climate change would affect the latitudinal, altitudinal and regional distribution of coffee production. As the GCM data used for this chapter differed from the data used in chapter 3 the impacts of climate change on the suitability distribution will be presented alongside the results of the distribution of harvested area as modeled by Globiom.

5.2.2.1 Latitudinal distribution

The distribution of suitability by latitude is shown in Figure 68a/b. Under current conditions the most suitable area for Arabica was confined to 27°S and 21°N, with large areas around 21°S and between 6°S and 9°N. Impacts were negative across all latitudes, time slices and GCMs. Differences between time slices and GCMs were limited to the degree of impact. An exception was the UKMO ("global dry") scenario for which a small positive impact at the extreme Southern latitude was projected. In the "global wet" MRI scenario at latitudes around 24°S suitability was higher in 2050 than in 2030 (Figure 68a).

Figure 68. Distribution of suitability and area by latitude



a/b Suitability distribution; c/d area modeled by Globiom; a/c Arabica; b/d Robusta (own data and representation).

The area suitable for Robusta showed similar margins as the Arabica area, from 24°S to 21°N, but with a much more pronounced concentration around the equator. Impacts were projected to be negative in general with some notable exceptions. In the CNRM scenario ("global mid-range") Southern latitudes between 3° and 21° South could gain in suitable area. The "wet" MRI scenario projected positive impacts on suitability across all latitudes by 2030, but negative impacts by 2050 (Figure 68b).

The distribution of area by latitude is shown in Figure 68c/d. The area data represents the Globiom output and was normalized to area shares to make the data comparable between scenarios (total area differed due to increasing future demand or climate change impacts). Changes compared to the current distribution differed between GCMs and time slices across all latitudes. For Arabica in the "no climate change" (NoCC) reference scenario a nearly equal latitudinal distribution of area was projected as under current conditions. All scenarios projected negative impacts around moderate latitudes of 9° South and about 10°N. In all GCM scenarios for 2050 positive relative changes were projected around the equator. All climate change scenarios projected small relative area increases north of 21°N, but only the UKMO scenario south of 21°S (Figure 68c). Projections of the future latitudinal distribution of Robusta area did not agree on common trends. Only towards extreme latitudes some positive relative change could be observed (Figure 68d).

The weighted mean absolute latitude of suitable area for Arabica under current conditions was 12°N/S. In line with the observations from Figure 68a this latitude was projected to remain unchanged under future conditions with the exception of the MRI scenario for which it was first reduced to 10° by 2030 and increased to 12° by 2050 (Table 20). For Robusta the weighted mean absolute latitude of suitable area was much lower at 6°. This latitude will increase across all GCM scenarios to 7°-9°.

Table 20. Weighted mean absolute latitude by species, scenario, time slice and modeling approach

Scenario		Arabica		Robusta	
Current	Suitability	12		6	
	Base area	13		7	
Future		**2030**	**2050**	**2030**	**2050**
No CC	with Globiom	13	13	8	9
CNRM	Suitability	11	11	8	9
	with Globiom	12	12	9	9
MRI	Suitability	10	12	7	8
	with Globiom	11	11	8	8
UKMO	Suitability	11	11	6	7
	with Globiom	13	13	7	7

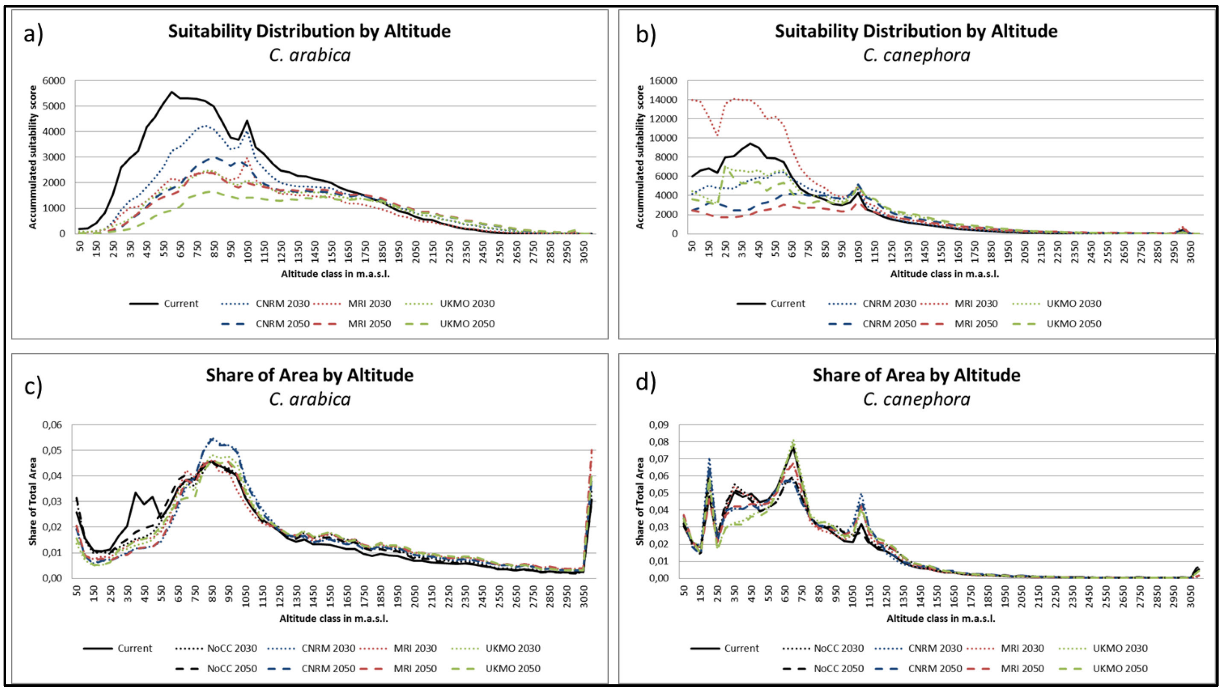
(Own data)

The share of area of Arabica as modeled by Globiom had an unchanged mean absolute latitude of 13° in the NoCC scenario. The CNRM and MRI scenarios projected this value to decrease to 11°-12°N/S, and the "dry" UKMO scenario to remain at 13°. Similarly, the Globiom output for Robusta area projected a slightly more pronounced migration to higher latitudes than the suitability shift (Table 20).

5.2.2.2 Altitudinal distribution

In Figure 69 the distribution of suitability and area is shown by altitude. For both species impacts were negative at low altitudes and positive at high altitudes. The direction of impacts shifted at about 1850m.a.s.l. for Arabica and 1050m.a.s.l. for Robusta. The exception was the "wet" MRI scenario which projected positive climate change impacts for Robusta by 2030 even at low altitudes, but equally highly negative impacts by 2050. The "mid-range" CNRM climate scenario projected a shift in direction for Robusta beyond ~800m.a.s.l. (Figure 69a/b).

Figure 69. Distribution of suitability and area by altitude



a/b Suitability distribution; c/d area modeled by Globiom; a/c Arabica; b/d Robusta (own data and representation).

Modeling the area with Globiom resulted in somewhat similar trends (Figure 69c/d). For Arabica all GCM scenarios, and the reference scenario with historical climate, projected relatively reduced area at low altitudes and relatively more area at higher altitudes. Depending on the scenario the change in direction of impacts was projected somewhere between 750m.a.s.l. and 900m.a.s.l. (Figure 69c). For Robusta the resulting future distribution of area shares was similar. Above approximately 550m.a.s.l. the impacts quality shifted from negative to positive, i.e. in future scenarios with climate change relatively more area will be cultivated at higher altitudes (Figure 69d).

The precise value of altitudes was not directly comparable between the suitability data and the area data because the resolution of Globiom is coarser than the suitability data. SimUs span ranges of altitude as was reflected by the inclusion of unfeasible altitudes below 300m.a.s.l. and beyond 3000m.a.s.l. (Figure 69c). Nevertheless, weighted mean altitudes were similar for the suitability data and the area data under baseline conditions (Table 21).

Table 21. Weighted mean altitude by species, scenario, time slice and modeling approach in m.a.s.l.

Scenario		Arabica		Robusta	
Current	Suitability	970		602	
	Base area	1000		681	
Future		**2030**	**2050**	**2030**	**2050**
No CC	with Globiom	1074	1042	701	694
CNRM	w/o Globiom	1115	1256	687	855
	with Globiom	1133	1135	689	708
MRI	w/o Globiom	1125	1300	556	908
	with Globiom	1135	1157	681	713
UKMO	w/o Globiom	1223	1398	725	785
	with Globiom	1154	1163	719	731

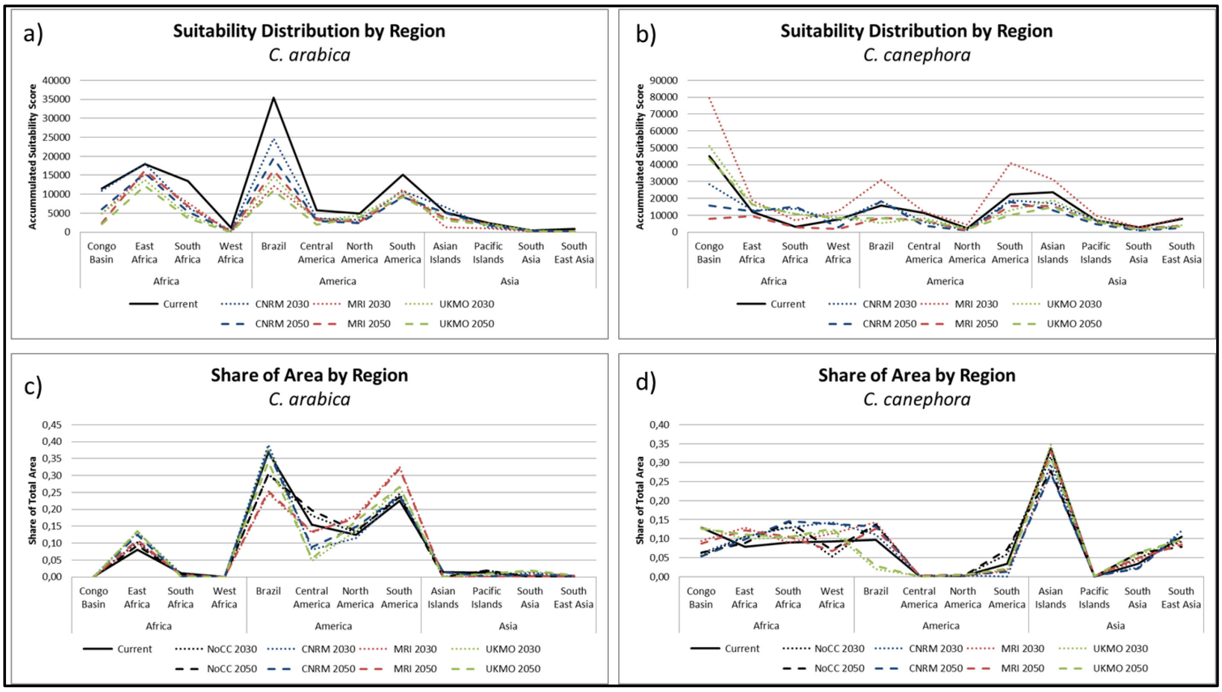
(Own data)

The current value of 602/681m.a.s.l. remained nearly unchanged for Robusta, but increased by about 50m from ~970/1000m.a.s.l. for Arabica, in the NoCC scenario. Climate change induced large shifts in mean altitude of suitability for Arabica between 300 and 400m until 2050. Robusta suitability will be 150m-300m higher on average by the same time (Table 21). Modeled with Globiom however, impacts were less pronounced. For Arabica the increase in altitude was seen to be around 100-150m, thereby higher than the increase in the baseline NoCC scenario. Mean altitude for Robusta area remained largely unchanged in the climate change scenarios. Only small shifts by 20-50m could be observed.

5.2.2.3 Regional distribution

The climate change impacts on suitability for Arabica were negative for all major coffee regions (Figure 70a). Impacts were mostly higher for 2050 than for 2030. Brazil was modeled to experience the highest absolute impacts, West Africa the highest relative impacts. The lowest relative impacts were projected for East Africa, and Asian and Pacific Islands. Figure 70b shows the distribution of suitability for Robusta coffee. By 2030 in most regions the impacts were potentially positive, especially in the "wet" MRI scenario. By 2050 however, impacts were negative in all regions, except East Africa and South Africa. Relative impacts for both time slices were highest for the South Asia and South East Asia regions where nearly 50% of suitability will be lost.

Figure 70. Distribution of suitability and area by region



a/b Suitability distribution; c/d area modeled by Globiom; a/c Arabica; b/d Robusta (own data and representation).

Again, the graph of the future distribution of area modeled by Globiom (Figure 70c/d) considered shares of total area. Therefore globally net impacts were zero. Across all scenarios, including NoCC reference scenarios, relative Arabica area increased substantially in East Africa, North and South America. Notable losers were Central America and Brazil with relative losses of up to 50% in most scenarios (Figure 70c). Robusta production was projected to diversify globally. In the baseline scenario a third of Robusta area was located in the Asian Island states, e.g. Indonesia and the Philippines and South East Asia. All climate change scenarios agreed that this region will lose relative shares, while East Africa, West Africa, Central America and the Pacific Islands win relative area (Figure 70d).

5.2.3 Spatial distribution of production

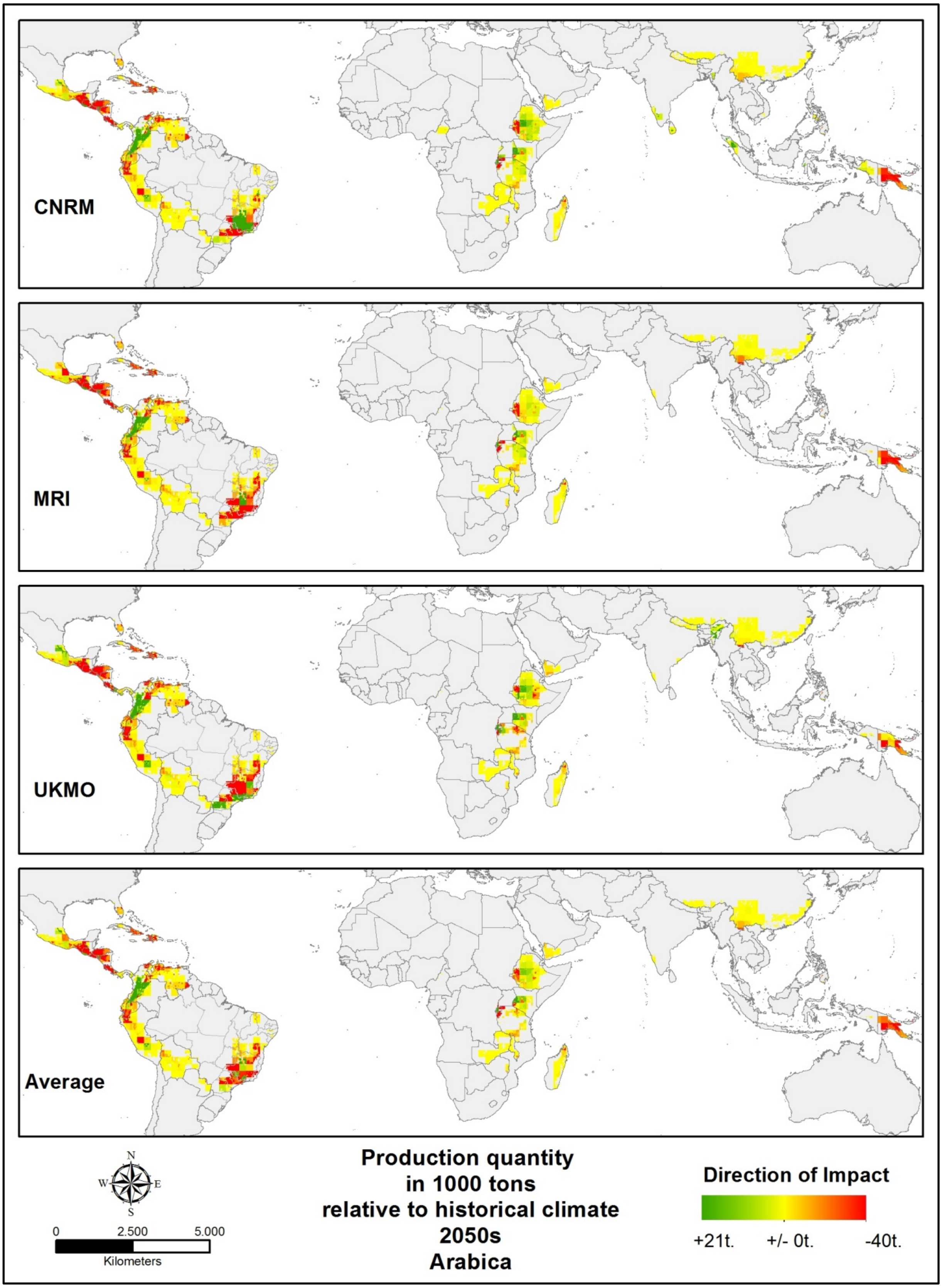
Globiom is spatially explicit. This allowed the mapping of the impacts of climate change on production. Figure 71 and Figure 72 show the distribution of production in the two coffee production systems by 2050, relative to the production in the reference scenario with historical climate.

Figure 71 and Figure 72 demonstrate the impact of climate change on the spatial distribution of coffee production. For each region net impacts were shown in Figure 70c/d. Here, the impact within regions is shown. Depending on the GCM scenario the change in the distribution may disagree substantially. E.g. for Arabica in Brazil and Robusta in India the GCM scenario changed the direction of impacts for some regions.

However, for some locations the scenarios agreed on impacts. In the Americas some Central American countries such as Nicaragua or El Salvador were found to be losers of Arabica production across scenarios, while Colombia will benefit. In East Africa the higher regions of Ethiopia will gain production, while the region towards Sudan loses. Kenya will produce more Arabica coffee in all scenarios (Figure 71). Impacts were relative to the NoCC scenario of the year 2050. As in all future scenarios most Arabica area in Asia was replaced by Robusta, no impacts on Arabica could be observed for this cropping system.

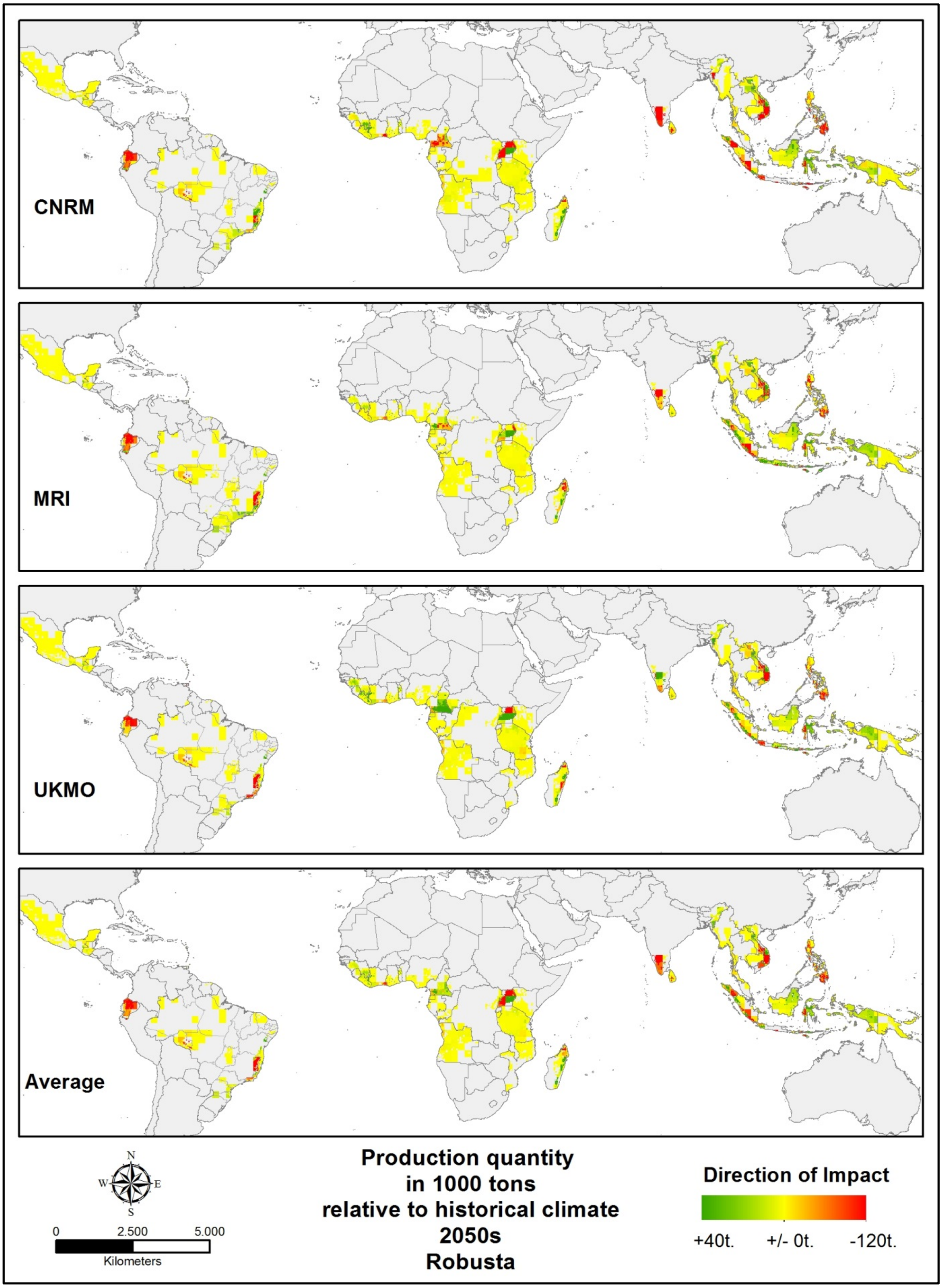
In Brazil Robusta production was projected to move north- or southwards from its current production region of Espirito Santo. In Africa Liberia will produce more coffee with climate change than with historical climate by 2050. Within Uganda Robusta production was shown to move towards the shores of Lake Victoria across scenarios, away from the interior. Vietnamese Robusta production moved northwards with its current locations in Dak Lak province losing substantially. In Indonesia and the Philippines the trend went towards higher altitudes and the interior of the islands (Figure 72).

Figure 71. Global distribution of produced Arabica quantities (2050) relative to the reference scenario



Produced Arabica quantity relative to reference scenario with historical climate in 1000 tons in the CNRM, MRI and UKMO scenarios and average impact by 2050 (own data and representation).

Figure 72. Global distribution of produced Robusta quantities (2050) relative to the reference scenario

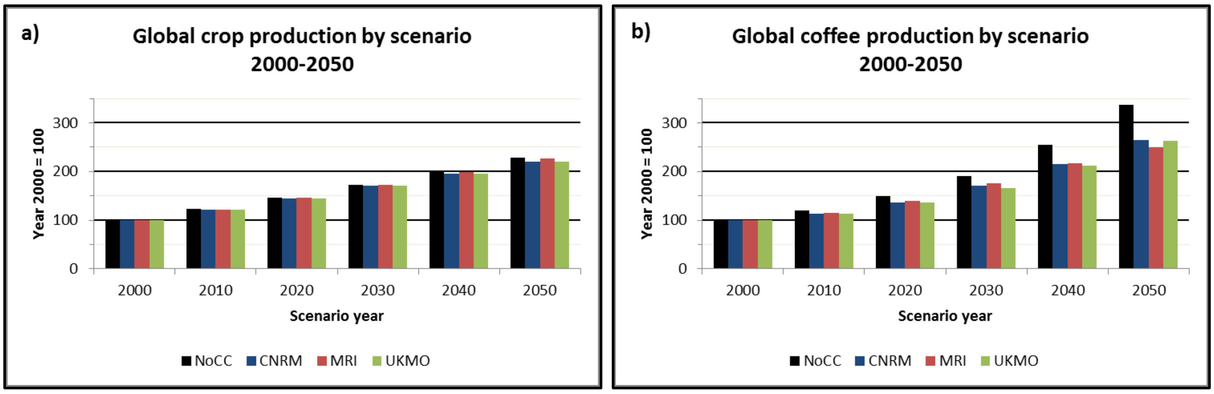


Produced Robusta quantity relative to reference scenario with historical climate in 1000 tons in the CNRM, MRI and UKMO scenarios and average impact by 2050 (own data and representation).

5.2.4 Total crop production and prices

As shown in Figure 67 the total production of coffee increased over time in all scenarios. In Figure 73 this development is put into context of the development of all other crop production.

Figure 73. Development of total crop production and coffee production

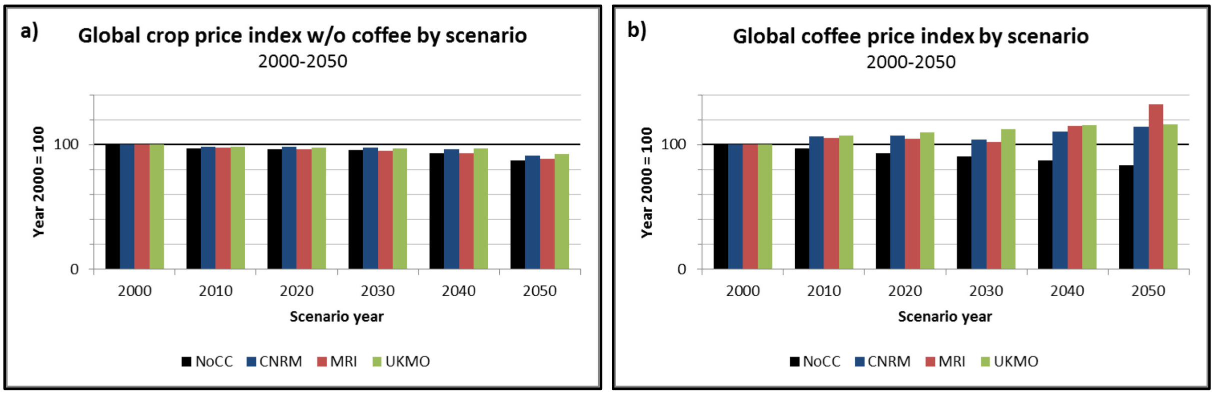


(a) Total crop production and (b) coffee production by GCM scenario (2000 – 2050) (own data and representation).

Relative to the base year 2000 global crop production increased by 128% in the scenario without climate change. With climate change this increase was lower. In the "wet" MRI scenario the total increase was 126%, in the "dry" UKMO scenario 121% and the "mid-range" scenario CNRM even lower at 120% until 2050. In the latter scenario global crop production was thus reduced by 4% compared to the NoCC scenario (Figure 73a). As described above this finding was in the same range as the results reported by (Mosnier et al. 2014) for calorie production. But impacts on coffee production were much more severe (Figure 73b).

Weighted mean prices for all crops and regions decreased until 2050 relative to the base year 2000 (Figure 74a). Without climate change prices decreased to 88% of the original level. The "wet" MRI scenario resulted in a similar price decrease to 89% of year 2000 levels. In the UKMO scenario prices are highest at 93%, and intermediate at 91% in the CNRM scenario. In the reference scenario without climate change coffee prices were reduced by 16% compared to current levels. This trend was reversed in the climate change scenarios. In the CNRM and UKMO scenarios the increase is 14-17%. In the MRI scenario by 2050 prices are 32% higher than in the base year and 58% higher than in the reference scenario (Figure 74b).

Figure 74. Development of crop price index and coffee price index



(a) Global crop price index and (b) global coffee price index by GCM scenario (own data and representation).

5.2.5 Scenario without coffee

In six model regions the base year revenue from coffee was at least 5% that of the revenue from the other 18 crops (Table 22). For these regions the effects on food security were analyzed by comparing the prices and total production of all other crops with and without coffee included in Globiom.

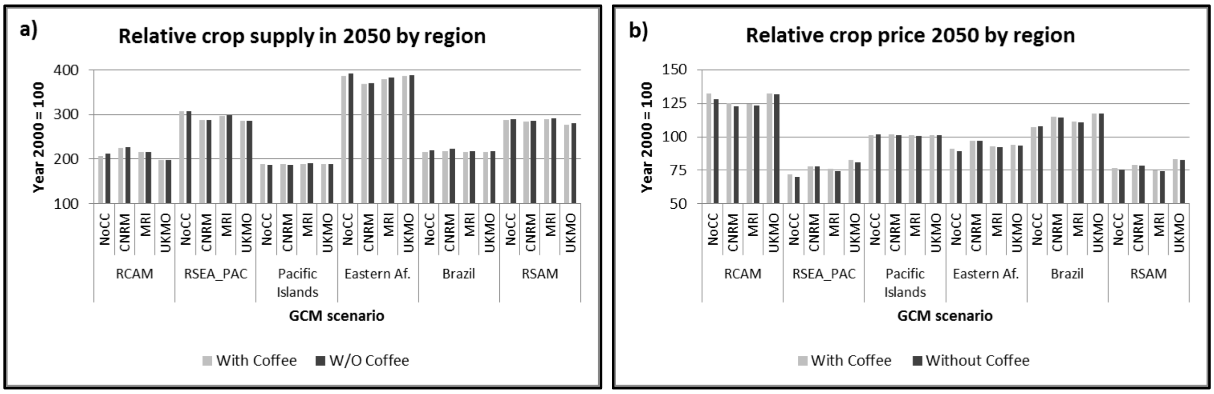
Table 22. Globiom regions with high revenues from coffee

Globiom Region	Countries	Coffee revenue[1]
RCAM	Bahamas, Belize, CostaRica, Cuba, DominicanRp, ElSalvador, Guadeloupe, Guatemala, Haiti, Honduras, Jamaica, Nicaragua, Panama, TrinidadTob	20%
RSEA_PAC	Cambodia, KoreaDPRp, Laos, Mongolia, VietNam	16%
Pacific_Islands	FijiIslands, FrPolynesia, NewCaledonia, PapuaNGuin, Samoa, SolomonIs, Vanuatu	9%
EasternAf	Burundi, Ethiopia, Kenya, Rwanda, Tanzania, Uganda	8%
BrazilReg	Brazil	7%
RSAM	Argentina, Bolivia, Chile, Colombia, Ecuador, FalklandIs, FrGuiana, Guyana, Paraguay, Peru, Suriname, Uruguay, Venezuela	6%

[1]Relative to revenue from other crops (own data)

Including coffee in Globiom was found to change the supply of other crops in these regions. In the scenario without coffee the supply of all other crops was higher than in the scenario without coffee. Figure 75a shows the relative supply of all crops, except coffee, for the regions identified above separated by GCM scenario in the year 2050. Across all regions and GCM scenarios the supply relative to the year 2000 increased more when coffee was not included. The average difference between the two scenarios for these regions was a 2% higher supply increase without coffee than with coffee. Prices on the other hand were lower without coffee included (Figure 75b). On average, prices were 1% lower without coffee included in the model.

Figure 75. Relative crop supply and relative crop prices by 2050



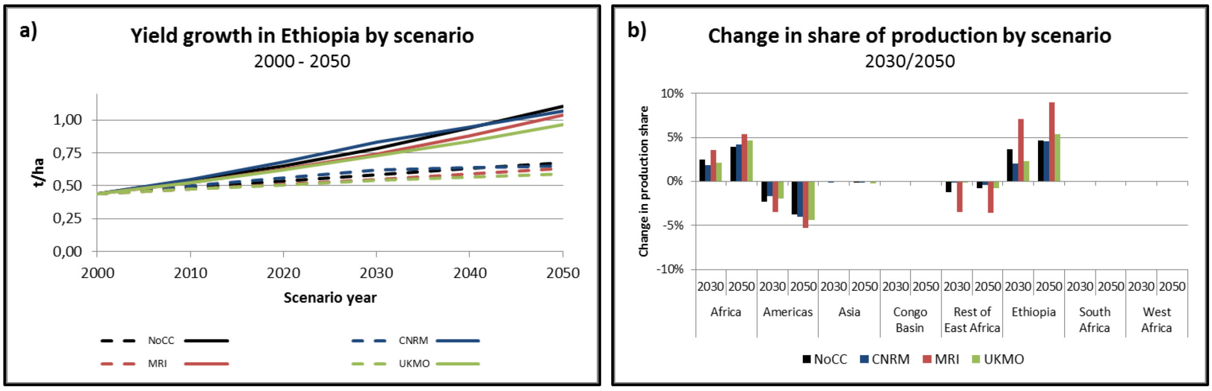
(a) Relative crop supply and (b) relative crop prices by 2050 in the scenarios with (light grey) and without coffee (dark grey) (own data and representation).

There were no feedbacks of climate change impacts on coffee on the supply of other crops. The impacts of climate change on crop supply relative to the scenario without climate change (NoCC) were similar in both scenarios. The average impact across all GCM scenarios on the crop supply in these regions was -1% for both the scenario that included coffee, and the scenario that excluded coffee. Similarly, in these regions prices were approximately 3% higher with climate change than without climate change in both scenarios. An exception from this general finding was the Central America region (RCAM). This region had the highest relative coffee production share in the base year 2000. Here, production was higher and prices lower in the climate change scenarios with coffee included, than without coffee included.

5.2.6 Opportunity scenario for Ethiopia

The scenario in which Ethiopia invests 15mill. USD per year in coffee research for an additional 1% yield increase is compared here with the scenarios presented above. As presented above (section 5.1.2.2) in the reference scenario the yield per hectare of area planted with *C. arabica* in Ethiopia increased because of technical progress from 440kg/ha to 670kg/ha by 2050 without climate change. With additional growth to close the yield gap this increased to 1110kg/ha by 2050. The yield was reduced in the GCM scenarios by 3% (CNRM), 6% (MRI), or 13% (UKMO) (Figure 76a).

Figure 76. Yield development in Ethiopia and change in production of Arabica coffee by scenario (2000-2050)



a) Yields (t/ha) in the reference scenario (dashed lines) and the added growth scenario (solid line); b) Production change of Arabica in Africa, the Americas and Asia and African regions by time slice and scenario in the scenario with added growth compared to the reference scenario (own data and representation).

This increased Arabica production in Ethiopia in all GCM scenarios and the scenario with historical climate compared to the scenario without added growth (Figure 76b). Without the policy Ethiopia's Arabica market share remained constant at 3% in the historical climate. With added growth but without climate change the market share increased to 7% by 2050. Without added growth, but with climate change, the market share of Ethiopia remained constant (MRI) or increased to 4-5% (CNRM, UKMO). With added growth, and climate change the market share by 2050 was 9-12%. The largest difference could be observed in the MRI scenario where the added growth raised the market share from 3% to 12%. Most of the additional shares of production were gained at the expense of the Americas as Africa as whole increased production shares. Only in the MRI scenario some of the market share gained came from other countries in Eastern Africa.

The net present value of the 15mill/year investment for the period 2000- 2050 was 715mill. for the scenario with climate change. In the CNRM and UKMO scenarios this was reduced to USD 303 and 341mill. respectively. In the MRI scenario the NPV was increased to USD 1.094 million. The investment had a positive NPV with discount rates as high as 20% in all scenarios (data not shown).

5.3 Discussion

Using a spatially explicit integrated climate change impact modeling framework it was shown that climate change will reduce future production by an amount that is approximately equivalent to the base period market volume. This reduction was driven by negative impacts on coffee yields especially on Arabica coffee which will lose disproportionally. The area necessary to meet future coffee demand will double in the future. Coffee prices will increase substantially unlike prices for other crops which tend to decrease. It was demonstrated that investments in coffee research would have a positive net present value for Ethiopia in all scenarios.

These results were achieved by integrating two production systems for coffee, Arabica and Robusta, in Globiom. This was done by employing the machine learning classification approach and the input data developed in chapters 3 and 4. Together with plausible demand side scenarios the resulting model added to the insights into the impacts of climate change on coffee production.

In the reference scenario with historical climate total production increased 3.5-fold from 7.3million tons to nearly 25million tons by the year 2050. This increase was reduced to less than 20million tons in the climate change scenarios. Additionally, with climate change world prices were up to 60% above reference scenario levels. This price increase appeared quite large, especially when compared to the price index for other crops that is a modest 3% above reference levels with climate change.

The reduction of demand in comparison with the reference scenario can be associated to the low price elasticities of demand for coffee. Globiom employs elasticities provided by the USDA (Seale et al. 2003) that were reduced to at least -0.3 from reported values in high income countries of up to -0.1. Also other literature reported values around -0.2 (Akiyama and Varangis 1990). In addition, elasticities in several regions were not adjusted with increasing GDP even when they were projected to transition from middle income to high-income. Taking this into account, the impact of climate change on the volume of the coffee market seems high, but the impact on prices was likely underestimated.

The demand scenario linked the SRES scenarios using per capita demand and GDP elasticities of demand with coffee consumption. Lacking a conclusive model in the literature to estimate changes of elasticities with income group the model here was a hybrid of historical patterns and evidence. Alternative demand scenarios could include the cultural contingency of coffee consumption. Great uncertainty stems e.g. from differences in coffee consumption between Asian economies. E.g. Japan has a comparatively low consumption, while in South Korea consumption has increased disproportionally with economic growth. Growth of coffee markets could therefore be well above the levels found here: In future periods major economies in South America, China, India and South East Asia transition to middle or high income economies. These economies could develop a high coffee demand like Northern countries or else a Japanese style tea culture.

Previous research had suggested a number of putative global trends for coffee production: latitudinal, altitudinal and regional migration (chapter 2.2). Climate change was found to have more severe impacts at low latitudes, and less so on higher latitudes. No positive effect of latitude on coffee suitability was found in the GCM scenarios. The comparison with area data modeled by Globiom showed that the latitudinal distribution of coffee production will essentially stay the same across GCMs and time slices. The negative impacts at low latitudes were not reflected in the future distribution of area.

Similarly, the clear altitudinal trend in the suitability data for both species and all scenarios was muted after Globiom modeled area distribution. Some altitudinal migration could still be observed, but especially at low altitudes not to the extent the suitability data suggested. A possible explanation for these effects is that production in Globiom was optimized for large grid cells (200km), while climate change often acts on smaller scales, e.g. a move in altitude by 200m. Given the relatively small scale of coffee production, in a 200km grid cell sufficient suitable area will remain to accommodate the coffee area. This will especially be true for lower altitudes where often vast areas, e.g. in Brazil or the Congo, were available. On the other hand, in high altitudes little agricultural area was available for efficient production.

Migration effects were much clearer between regions. For Arabica East Africa was found to be a relative beneficiary of climate change, as well as South America in comparison to Brazil. In the case of Robusta the effect was not as pronounced, even though African regions were found to be relative beneficiaries at the expense of South American Robusta production.

However, changes in on suitable area were higher than the resulting changes in production. Regions with high coffee production such as Brazil or Indonesia will have comparative advantages beyond climatic suitability. At the same time, other regions with low productivity e.g. in Africa will not be able to seize from the favorable climatic conditions unless productivity is raised. Such effects were reflected in Globiom in the calibrated cost parameter. Here, the cost parameter was assumed to remain constant over time. Any change in yield potential due to climate change must therefore have stronger effects than effects from other comparative advantages. The Brazil region was thus found to retain some of its market share in future decades despite the severe climate change impacts on suitability because of its advantages in productivity.

Nevertheless, remaining competitive will be a challenging task as exemplified by the sub-regional spatial effects of climate change. For some regions GCM scenario results disagreed on the spatial distribution of climate change impacts. E.g. in Brazil the global market share remained high, but each GCM scenario allocated the area required to keep up the production into different grid cells. However, the yield potential model for coffee assumed perfect adaptation. How such perfect adaptation should work remains unclear when considering the multi-decadal investment depth of coffee plantations and the simultaneous lack of reliable climate projections to ex-ante evaluate possible adaptation strategies. The regional effect of climate change will therefore be more pronounced than modeled here.

One aspect of future coffee production will be whether the Arabica production may compete with Robusta production in the future. The comparison of suitability data did not give conclusive insights: suitable area did not equate harvested area. The results obtained showed that the share of Arabica production on the world market will decrease further to about 50%. This was driven by the higher impact of climate change on Arabica yield than on Robusta yield. Here, Arabica and Robusta were seen as perfect substitutes on a singular coffee market. Previous research had shown that such an assumption is reasonable (Ghoshray 2010). Yet, the retail market is increasingly differentiated (Lewin et al. 2004) so that the assumption of perfect substitutability can be deemed to be challenged in the future. On global scale the magnitude of the decrease of higher quality Arabica share did not appear to be substantial enough to invalidate the perfect substitutes assumption. In addition, growth on coffee markets often comes from novel markets that have a preference for Robusta based convenience products (Lewin et al. 2004).

However, on regional scale a differentiation of demand within Globiom would likely change model outcomes. E.g. in Asia *C. arabica* based production was projected to vanish and to be replaced by *C. canephora*. In South America it was the other way round: *C. canephora* largely vanished. From an economic perspective this can be explained with the comparative advantage that each of the two crops hold in the two regions. The model therefore iteratively allocated area to the more efficient system. This would result in a situation where all demand in Asia is covered by Robusta coffee. This, however, is unfeasible given the existing demand for high quality coffee in the region. Including an additional condition in Globiom, e.g. one that demands that the coffee mix in all regions equals the global mix, would therefore likely result in additional Arabica area in Asia, and more Robusta area in South America, unless trade costs become much lower than the differences in production costs.

For the global coffee market the switch from Arabica to Robusta is likely the most important means of adaption, which is why this production systems differentiation was made here. Nevertheless, another important aspect is the use of shade to adapt to unfavorable climatic conditions, the use of fertilizer to increase productivity, and irrigation that achieves both goals. Ideally the supply side model would be able to differentiate these systems for coffee, like Globiom already does for other crops. In chapter 4 the basis was set to differentiate shaded and unshaded production as well. However, more research will be needed to model impacts of climate change on yield potential for the different systems.

In addition to better production system differentiation, a better production system definition will be useful. Including more data for each system would likely change the direction of impacts of climate change in some regions. The coffee model lacked the integration of water demand restrictions. This has been mentioned as a concern for South East Asian Robusta locations (D'haeze et al. 2005) and Central American Arabica production where even today groundwater depletion and excessive water contamination caused by coffee production is a concern. Additionally data on carbon stocks and carbon emissions should be integrated in order to enable the evaluation of trade-offs between adaptation and mitigation strategies.

The solution to estimate the yield potential from suitability data from chapter 3, was both pragmatic and clearly reproducible across time slices. Estimating the relationship suitability-yield potential based on real data was not feasible due to a lack of reliable data. Large scale crop simulation models are often calibrated with observed yield data from Monfreda et al. (2008) (Rosenzweig et al. 2014) which was found unfit for coffee purposes (Eriyagama et al. 2014). In exploratory attempts only low correlations between such observed yield patterns and suitability were found (data not shown). However, for example the widely recognized FAO AEZ model applies yield penalties to maximum yield potential when optimality conditions for a crop at a site are not fulfilled (Fischer et al. 2012). Also, in the case of maize in South Africa a methodologically similar suitability model was able to reproduce yield patterns with comparable certainty as crop simulation models (Estes et al. 2013). The approach chosen here was therefore in line with comparable studies.

The assumed biophysical impacts on coffee were comparable with other crops in Globiom. Climate change reduced yield increases from technical progress to about nil in the case of

Arabica, and 20-25% for Robusta. Globally, this impact was harsher than the impact on other crops, which were projected to produce yields about 3% below reference levels. However, the EPIC model assumed perfect adaptation and CO_2 fertilization effects (Mosnier et al. 2014). In a model intercomparison study that compared the results of several similar studies found mean biophysical impacts of about -17% (Nelson et al. 2014), but impacts may be higher with up to -38% (Müller and Robertson 2014).

Another possible explanation for the different magnitude of climate change impacts is that most other crops in Globiom were adapted to a much wider range of climatic conditions throughout history than coffee. I.e. maize is produced in both tropical and temperate climates. Therefore the adaptive capacity of this crop can be assumed to be higher than for coffee which is only cultivated within a very narrow agro-ecological niche. Interestingly, the one other crop that was projected to experience impacts of similar magnitude as coffee was sugar cane. This crop grows in similar regions as Robusta coffee, the hot and humid tropics. Most research focusses on global staples. It will be interesting to see how other tropical crops will perform in future climate.

Without changes in the coffee market in future periods, the production of other crops was higher, and prices were lower. And, this impact of climate change on the other crops relative to the scenario with historical climate was unchanged with the inclusion of coffee (with one exception). These findings provide support for the argument that increased food security may only be achieved if also demand side issues are considered (Bajzelj et al. 2014).

The inclusion of rising coffee demand reduced crop production by 2% in regions where coffee represents a substantial share of agricultural production. This effect of increasing coffee demand on other crops is of the same magnitude as climate change (Mosnier et al. 2014). This finding can be easily explained with the area available in this scenario that would otherwise be used for coffee production. In Central America where the value of coffee production was as high as 20% of the value of other crop production, in the GCM scenarios prices were found to be reduced and production of other crops increased. This suggested that area that without climate change would be occupied by coffee was allocated to production of other crops in Central America. No opposite production effects in other regions were found. The coffee area lost in Central America was thus replaced by area in other regions without affecting production of other crops.

However, in Central America coffee provides large shares of the export revenues (Figure 26, (ICO 2014)), and employs substantial shares of the rural workforce. The loss of income from coffee could therefore have negative poverty effects that offsets the effects from improved accessibility of food. Within the partial equilibrium model Globiom this cannot be evaluated. The use of a CGE model of the region that also differentiates different labor classes would give additional insight into the consequences of the loss of coffee income in Central America.

As a simple example application of Globiom with coffee included it was shown that for Ethiopia an investment in coffee research would always have a positive net present value, but that the return varied substantially between the GCM scenarios. The rationale of the policy was that climatic suitability for coffee in Eastern Africa will be less impacted by climate change than in other regions. The revenue from the added growth in productivity was meant to be leveraged by the negative impacts of climate change elsewhere.

Against the expectation, climate change had a positive leverage effect on the investment only in the MRI scenario. In two of the three climate change scenarios (CNRM and UKMO) the net present value was found to be reduced. The reason was that with increasing heat Robusta will become relatively more competitive in Eastern Africa. Arabica will struggle to compete even when becoming highly productive, reducing returns from Arabica production. Using Globiom to model trade-offs between the two species will therefore provide additional insight also in future applications. However, differentiating the demand side would likely change the investment outcome of the policy. Arabica commands a price premium; the competition with Robusta is not as perfect as modelled here. As Arabica coffee is more severely impacted by climate change than Robusta the price gap will increase. In this case the additional revenue from Arabica would be higher than modeled here, and the policy would consequently result in higher net present values with climate change.

5.4 Conclusions

In this chapter the integration of a coffee model in the partial equilibrium model Globiom was demonstrated. The input data was based on the methods developed in chapters 3 and 4, or was defined in coherence with the model set up. This chapter concluded the modelling chain from emission scenarios, global climate models, biophysical impact model and partial equilibrium model. The result was an integrated impact assessment of climate change on global coffee production.

In this integrated assessment of the climate change impacts on global coffee production a comparatively strong impact was found. Total production was reduced by 25% and prices 60% higher than with historical climate. An interesting question for future research that could not be resolved here will be whether the difference in impacts was caused by the differences in the biophysical impact model, or because the character of the coffee crop is fundamentally different.

In comparison with the suitability model from chapter 3, the impacts were found to be somewhat muted after inclusion of Gobiom. Latitudinal effects, altitudinal impacts and also regional impacts were more pronounced with the pure suitability model. Two aspects were argued to cause this: comparative advantages of current production regions over novel regions, and the coarse resolution of Globiom.

From a development and climate adaptation perspective the former aspect will be of high interest: To what extent will comparative advantages stemming from historical periods, such as infrastructure, management knowledge etc. be transferable to future periods? Adaptation to climate change can be achieved by knowledge transfer. In this case a region like Brazil with a well-developed research and knowledge infrastructure in the coffee sector would be able to seize from its existing comparative advantage. On the other hand, several regions in Eastern Africa or Asia will be much less impacted by climatic changes. In these cases technological jumps appear possible with much less effort. An example of such a policy was modeled for Arabica production in Ethiopia. The results from this experiment demonstrated that the use of Globiom will be a useful tool to evaluate such policies.

6 Discussion and Conclusions

Climate change and increasing demand will be the paradigms that shape the coffee sector in the future. In this thesis I showed that by 2050 on only half the area that is currently available for coffee production 2.5-times as much coffee will have to be produced to meet future demand. Reduced yields and increased prices were shown to reduce the coffee market by more than 5million tons per year, equivalent to the size of the baseyear 2000 market volume. An integrated modeling framework was developed throughout the thesis to combine both supply side impacts of climate change on coffee production and demand side changes on consumption due to population growth and increasing global income.

Previously, two and a half gaps existed in the literature on coffee. No globally coherent biophysical impact study for both crops existed even though regional studies suggested drastic impacts (chapter 2.2). No partial equilibrium model had been used to study the implications of probable demand increases (chapter 2.3) for coffee production in future periods. Existing data on the physical distribution of coffee production was not useful for the intended use here: the integration of biophysical impacts with demand side effects (Eriyagama et al. 2014). In this thesis these knowledge gaps where addressed by the development of novel methodological approaches. First, machine learning classification was used for a biophysical impacts model (chapter 3). The resulting data of climatic suitability led to the development of a substantially improved dataset of the spatial distribution of coffee production (chapter 4). These two steps were preconditions to include a model of the coffee sector in the spatially explicit partial equilibrium modeling framework Globiom (chapter 5).

The novelty of the work presented here therefore lies not only in the results, but also in the method. A common theme throughout this thesis was that the conventional approaches which can be used for other crops to close the aforementioned knowledge gaps were not applicable for coffee. Most research to investigate the climate change impacts on crop production uses crop simulation models and is focused on a very limited set of crops (White et al. 2011). Hence, before this work, not only was there no integrated climate change impact assessment of the coffee sector, nor did the necessary tools to conduct such research exist.

The advantages of the conventional, simulation model based, approach are emphasized whenever such research is published: all aspects of the crop management may be modeled in a single model. Crop growth and crop management are simulated to derive information on yield potential, resource use and production costs. Just as often do authors discuss the limitations to these models: Model ensembles are able to reproduce historic yields, individual models are not (Rotter et al. 2011). Simulation models are complex; understanding model differences is subject to large research projects (Rosenzweig et al. 2014). The single crop simulation model for coffee Caf2007 (Van Oijen et al. 2010b) was therefore unlikely to provide reasonable results to match the knowledge for large staple crops, especially when taking into account technical limitations. A variety of alternative approaches has therefore been used for coffee before this thesis. On global scale applicable were simple envelope approaches based on a limited set of variables (e.g. Zullo et al. 2011) or machine learning approaches (Läderach, Lundy, et al. 2011). The latter, however, a) have been shown to be prone to overestimate impacts of climate change; b) had not been applied on global scale for coffee.

The first innovation of this work was to overcome these limitations by developing a global impacts model using a machine learning ensemble approach. This was the first published research to assess the climate change impacts on both Arabica and Robusta coffee on global scale in a coherent way (Bunn et al. 2015). Furthermore, the methodology that was applied was an improvement when compared to previous research on the regional impacts. Special emphasis was put on model generalization both by carefully choosing algorithm parameters, but also by using an ensemble approach that reflects model uncertainty. This was meant to avoid an over estimation of climate change impacts in novel climates caused by a too close fit of models to known climatic conditions. No comparable historic reference data was available to test the ability of the approach here to correctly project across time slices. Nevertheless, the validity of the approach here was demonstrated by two indications: even when trained on highly clustered occurrence data the models used here correctly predicted spatially distant occurrences. By 2030 in the MRI climate scenario positive impacts on Robusta production were found.

The advantage of the approach here was its limited need for data. It could therefore be readily applied to crops where the limitations to impacts modeling are comparable (e.g. cocoa). Using geo-references of occurrence locations and climate data a reasonable model of the spatial distribution of coffee production could be developed. This however, came at the cost of accepting some disadvantages when compared to simulation model approaches. The latter are capable of explicitly modeling resource and input use and adaptation management (Rosenzweig et al. 2014). The machine learning classification approach rates the climate independent of management options or resource availability. The importance of adaptation options such as shaded coffee production, irrigation use or novel varieties could therefore not be assessed here. A solution to this problem would be to differentiate not only between suitable and unsuitable climates but several agro-climatic zones (Bunn et al. Submitted).

The second innovation of this thesis addressed the lack of spatially explicit data of coffee production. Most research that requires knowledge on the spatial distribution of area used for crop production uses either data from MapSpam (You et al. 2012) or Monfreda (Monfreda et al. 2008). This data was previously shown to be of limited use value for coffee (Eriyagama et al. 2014). The data does not differentiate by coffee species, nor does it reflect the distribution of coffee production correctly. Here, a first step was done towards a better understanding of the physical distribution of coffee production by generating data that differentiated between the two major coffee species and the two most important production systems. The distribution data from the previous chapter provided a good approximation of the true distribution of coffee production as estimated by independent production data. Especially for regions where no subnational data was available as input to the previous datasets, the new data approximated much better the observed distribution.

Additional research merits the question how the use of machine learning classification data could support the improvement of global gridded crop simulation models. Such models use disaggregated production statistics for model calibration and validation (Rosenzweig et al. 2014). The datasets by Monfreda et al. (2008) and You and Wood (2006) present crucial shortcomings for this use. Where they do not use suitability data in some cases production is allocated to locations that are unfeasible because of their climatic characteristics (e.g. in China). But, the use of the GAEZ crop yield potential model (Fischer et al. 2012) to avoid such infeasibilities is not without problems either. When the resulting data is used to validate the ability of simulation models to reproduce observed yields, this model fit would be compared against data that includes a yield estimation model itself. The use of machine learning ensemble derived suitability data as an input to estimate the spatial extent of crop production to feasible locations should be considered as an alternative approach. The method developed here estimated the spatial distribution with high certainty, without being a crop yield model itself.

Some of the steps for the disaggregation here were specific to coffee, such as allowing production in locations that are forest covered. Nevertheless, existing datasets for crops that are not key staple crops are often equally uncertain as the data for coffee. Such data could be easily improved using the method developed here. Research should however add a differentiation into irrigated production and rain-fed to facilitate research concerning water scarcity.

The explicit inclusion of adaptation and resource variables in the modeling process could alter some of the results presented in the last chapter. E.g. irrigation of coffee production is already a limitation to coffee production in several regions (Eriyagama et al. 2014). In coming decades coffee harvested area was projected to double but with less area suitable for coffee cultivation. Therefore production will be increasingly agglomerated in a few regions. Locally, this will increase the pressure on resource endowments such as ground water reserves. An inclusion of resource constraints for coffee in Globiom could therefore alter the model outcomes as reported here. This will be of additional importance when irrigation becomes an increasingly important means to raise productivity and to adapt to heat and drought. Similarly, the use of biodiverse shade will be an important means of adaptation, and will also have implications for mitigation efforts.

Overcoming the limitations that previously hindered the integration of coffee data in economic-biophysical integrated modeling frameworks permitted the evaluation of the effects of climate change and associated demand increases. Previous research emphasized the need to regard climate change impacts in global partial equilibrium context (Bajzelj et al. 2014). Biophysical impacts of climate change on crop yields were found to be partially offset by changing economic incentives (Nelson et al. 2014). Model intercomparison found mean biophysical shocks of 17% yield reduction. However, the total impact on consumption was a reduction of 3% due to economic incentives that resulted in shifts to production systems with higher yields, shifts in trade, increased area and changed consumption patterns (Nelson et al. 2014). Here, for the first time a similar effect for coffee was shown. Suitable area was found to be reduced by about 50% for both crops. This translated into yield reductions of about 34% for Arabica and 17% for Robusta relative to historic climate. After partial equilibrium analysis total consumption was found to be reduced by 22-29%, while total area for coffee production was similar in all scenarios. Even though the biophysical impact was reduced by economic incentives impacts on coffee were thus considerably higher than on other crops. This was explained with the very specific climatic requirements of the coffee crop which limits adaptation by migration or systems change.

This thesis contributed to the understanding of limits to adaptation in the coffee sector. Previously, latitudinal migration, altitudinal migration of production, or replacement of Arabica with Robusta had been suggested to respond to climate change. These studies were largely based on a few variables or did not seek to explain the projected trends by specific bioclimatic effects (chapter 2.2). The machine learning approach that was used here showed that for Arabica heat stress determines the spatial distribution. Robusta was largely confined to locations with an even climate without seasonal and diurnal temperature variation. These variables were found to rule out a general latitudinal migration of coffee production, but a global trend that production will migrate in elevation was confirmed.

The adaptation means that is most widely discussed in the coffee sector is the replacement of Arabica with Robusta production. Here, only a partial answer could be provided. The differentiation of demand by crop in Globiom would provide further insight into the shift between the two coffee systems. Both products were shown to function as a single market; price changes for one crop altered price for the other crop with some delay (Ghoshray 2010). However, saturated markets in high income countries grow not by volume but by disproportional demand for high quality Arabica (Lewin et al. 2004). It remains to be investigated whether in coming decades demand for Arabica creates the necessary price incentives to balance the relatively worse biophysical impacts on yields.

Prices for coffee were found to be increased by 60% until 2050 compared to the scenario with historical climate. This price increase was high compared to other crops which generally see price decreases (Mosnier et al. 2014), modest increases (Fischer et al. 2005) or increases of 20% (Nelson et al. 2014). A limitation to this study was the consideration of mean effects without including variability. Price fluctuations of the magnitude found here are not uncommon on coffee markets. Because of the very small short term elasticities of both supply and demand price uncertainty is high for both consumers and producers.

With increasingly uncertain climatic conditions such price fluctuations could be exacerbated. For example Hansen, Sato and Ruedy (2012) stated that climate change means the increasing likelihood of a (locally) very unusual climate event to occur, e.g. a heat spell once in 10 years rather than once in 100 years. It is this kind of climate variability that results in reduced yields (Gay Garcia et al. 2006; Craparo et al. 2015). The resulting uncertainty will affect stakeholders along the entire supply chain. The high economic risk has been found to be a major reason to drive producers towards more reliable income sources (Baca et al. 2014). Fluctuations on coffee markets will increase to an extent that even global trade houses will not be able to offset the risk by using regionally diversified portfolios due to climate change (Joann de Zegher, personal communication, forthcoming). Beyond the mean effect reported here, the impact of climate change on coffee markets could thus be even higher. Despite the challenges to the development of useful crop simulation models that were discussed here, the availability of such a model would improve insight into this issue.

Coming research that focusses on climate change impacts should focus on the identification of unambiguous change trajectories to facilitate the development of feasible adaptation strategies. In this thesis modeling uncertainty was largely left aside to be able to demonstrate the modeling framework. The hierarchical framework used here passed mean results to subsequent modeling steps (Figure 1). This was despite the considerable modeling uncertainty that adds during each step of the framework: emission scenarios, GCMs, crop model, data disaggregation and integrated demand modeling. E.g. chapter 3 results were reported as "a 50% reduction in available area", although the mean global effect for the RCP 2.6 scenario was a 40% reduction and 60% for RCP 8.5 (Table 13). Considering the most optimistic impact across all GCMs in RCP 2.6 resulted in a mere 18% reduction, the worst outcome of the RCP 8.5 would be a 80% reduction of available area. The evaluation metrics for the classification models showed a good prediction capability of test sites (Table 11). Yet, the resulting distribution of each model varied substantially (data not shown). As the suitability data was a crucial input for the spatial disaggregation also the chapter 4 data represented only one of several feasible spatial distributions. In the economic modeling step in chapter 5 a single SRES trajectory was considered for three of many GCMs. As discussed, the demand scenario was also highly uncertain because of the cultural contingency of coffee consumption. Nevertheless robust results could be derived in this modeling framework.

Throughout this thesis it was demonstrated that climate change will have a profound negative impact on global coffee production, independent of emission scenario, climate or crop model. The bad news is the increasing heat which will reduce yields, make area unsuitable for production, and water management tougher. On the plus side, the sector as a whole will be presented with novel opportunities from increasing demand. Adaptation to climate change will be a major challenge for producer countries, especially given the considerable uncertainty in climate modeling on local scale. But for those who manage to find smart solutions higher prices will offer attractive rewards. Thus, there will be coffee on the table in 2050, but it will be of lower quality, will cost more and it will still be in the focus of sustainable enterprises because its production will still be shaped by poverty risk and environmental problems.

7 Literature

Akiyama, T., and P.N. Varangis. 1990. “The impact of the International Coffee Agreement on producing countries.” *The World Bank Economic Review* 4(2):157–173.

Alexandratos, N., J. Bruinsma, G. Bödeker, J. Schmidhuber, S. Broca, P. Shetty, and M. Ottaviani. 2006. “World agriculture: towards 2030/2050.” Prospects for food, nutrition, agriculture and major commodity groups Food and Agriculture Organization of the United Nations. Available at: http://www.fao.org/docrep/009/a0607e/a0607e00.htm.

Anderson, R.P., and I. Gonzalez. 2011. “Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent.” *Ecological Modelling* 222(15):2796–2811.

Anderson, W., L. You, S. Wood, U. Wood-Sichra, and W. Wu. 2015. “An analysis of methodological and spatial differences in global cropping systems models and maps.” *Global Ecology and Biogeography* 24(2):180–191.

Anonymous. 1913. “The Production of Coffee in Brazil.” *Journal of the Royal Society of Arts* 61(3146):449–451.

Araujo, M., and M. New. 2007. “Ensemble forecasting of species distributions.” *Trends in Ecology & Evolution* 22(1):42–47.

Assad, E.D., and H.S. Pinto. 2008. “Global Warming and the New Geography of Agricultural Production in Brazil.”

Assad, E.D., H.S. Pinto, J. Zullo Jr, and A.M.H. Ávila. 2004. “Impacto das mudanças climáticas no zoneamento agroclimático do café no Brasil.” *Pesquisa Agropecuária Brasileira* 39(11):6–11.

Baca, M., P. Läderach, J. Haggar, G. Schroth, and O. Ovalle. 2014. “An Integrated Framework for Assessing Vulnerability to Climate Change and Developing Adaptation Strategies for Coffee Growing Families in Mesoamerica” B. Bond-Lamberty, ed. *PLoS ONE* 9(2):e88463.

Bajzelj, B., K.S. Richards, J.M. Allwood, P. Smith, J.S. Dennis, E. Curmi, and C.A. Gilligan. 2014. “Importance of food-demand management for climate mitigation.” *Nature Clim. Change* 4(10):924–929.

Barbet-Massin, M., F. Jiguet, C.H. Albert, and W. Thuiller. 2012. “Selecting pseudo-absences for species distribution models: how, where and how many?” *Methods in Ecology and Evolution* 3(2):327–338.

De Beenhouwer, M., R. Aerts, and O. Honnay. 2013. “A global meta-analysis of the biodiversity and ecosystem service benefits of coffee and cacao agroforestry.” *Agriculture, Ecosystems & Environment* 175:1–7.

Bhatt, S., P.W. Gething, O.J. Brady, J.P. Messina, A.W. Farlow, C.L. Moyes, J.M. Drake, J.S. Brownstein, A.G. Hoen, O. Sankoh, M.F. Myers, D.B. George, T. Jaenisch, G.R.W.

Wint, C.P. Simmons, T.W. Scott, J.J. Farrar, and S.I. Hay. 2013. "The global distribution and burden of dengue." *Nature* 496(7446):504–507.

Bosselmann, A.S. 2012. "Mediating factors of land use change among coffee farmers in a biological corridor." *Ecological Economics* 80:79–88.

Boulesteix, A.-L., S. Janitza, J. Kruppa, and I.R. König. 2012. "Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2(6):493–507.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45(1):5–32.

Bunn, C., P. Läderach, O. Ovalle Rivera, and D. Kirschke. 2015. "A bitter cup: climate change profile of global production of Arabica and Robusta coffee." *Climatic Change* 129(1-2):89–101.

Bunn, C., P. Läderach, J.G. Perez, C. Montagnon, and T. Schilling. Submitted. "Multiclass classification of agro-ecological zones for arabica coffee: An improved understanding of the impacts of climate change." *PLOS ONE*.

Clarke, L., J. Edmonds, H. Jacoby, H. Pitcher, J. Reilly, and R. Richels. 2007. "Scenarios of greenhouse gas emissions and atmospheric concentrations." *US Department of Energy Publications*:6.

Craparo, A.C.W., P.J.A. Van Asten, P. Läderach, L.T.P. Jassogne, and S.W. Grab. 2015. "Coffea arabica yields decline in Tanzania due to climate change: Global implications." *Agricultural and Forest Meteorology* 207:1–10.

Crawford, J. 1852. "History of Coffee." *Journal of the Statistical Society of London* 15(1):50–58.

Davis, A.P., T.W. Gole, S. Baena, and J. Moat. 2012. "The Impact of Climate Change on Indigenous Arabica Coffee (Coffea arabica): Predicting Future Trends and Identifying Priorities" B. Fenton, ed. *PLoS ONE* 7(11):e47981.

Descroix, F., and J. Snoeck. 2009. "Environmental Factors Suitable for Coffee Production." In *Coffee: Growing, Processing, Sustainable Production*. Weinheim: Wiley-VCH, pp. 168–181.

D'haeze, D., J. Deckers, D. Raes, T.A. Phong, and H.V. Loi. 2005. "Environmental and socio-economic impacts of institutional reforms on the agricultural sector of Vietnam: Land suitability assessment for Robusta coffee in the Dak Gan region." *Agriculture, Ecosystems & Environment* 105(1–2):59–76.

Diniz-Filho, J.A.F., L. Mauricio Bini, T. Fernando Rangel, R.D. Loyola, C. Hof, D. NoguÃ©s-Bravo, and M.B. AraÃºjo. 2009. "Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change." *Ecography* 32(6):897–906.

Dormann, C.F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J.R.G. Marquéz, B. Gruber, B. Lafourcade, and P.J. Leitão. 2013. "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance." *Ecography* 36(1):027–046.

Dudík, M., S.J. Phillips, and R.E. Schapire. 2005. "Correcting sample selection bias in maximum entropy density estimation." In *Advances in neural information processing systems*. pp. 323–330.

Dumollard, G., P. Havlík, and M. Herrero. 2013. "Climate change, agriculture and food security: a comparative review of global modelling approaches."

Eakin, H., A. Winkels, and J. Sendzimir. 2009. "Nested vulnerability: exploring cross-scale linkages and vulnerability teleconnections in Mexican and Vietnamese coffee systems." *Environmental Science & Policy* 12(4):398–412.

Elith, J., and C.H. Graham. 2009. "Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models." *Ecography* 32(1):66–77.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. Lohmann, B. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, and N.E. Zimmermann. 2006. "Novel methods improve prediction of species' distributions from occurrence data." *Ecography* 29(2):129–151.

Elith, J., S.J. Phillips, T. Hastie, M. Dudík, Y.E. Chee, and C.J. Yates. 2011. "A statistical explanation of MaxEnt for ecologists." *Diversity and Distributions* 17(1):43–57.

Eriyagama, N., Y. Chemin, and R. Alankara. 2014. "A methodology for quantifying global consumptive water use of coffee for sustainable production under conditions of climate change." *Journal of Water and Climate Change* 5(2):128.

Estes, L.D., B.A. Bradley, H. Beukes, D.G. Hole, M. Lau, M.G. Oppenheimer, R. Schulze, M.A. Tadross, and W.R. Turner. 2013. "Comparing mechanistic and empirical model projections of crop suitability and productivity: implications for ecological forecasting." *Global Ecology and Biogeography*:n/a–n/a.

European Commission. 2003. "Global Land Cover 2000 database." Available at: http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php [Accessed December 21, 2012].

FAO. 2014a. "Agro Maps." Available at: http://kids.fao.org/agromaps/ [Accessed May 19, 2014].

FAO. 2011. "Ecocrop." *Food and Agriculture Organization of the United Nations*. Available at: http://ecocrop.fao.org/ecocrop/srv/en/home [Accessed May 13, 2011].

FAO. 2012. "FAOSTAT." Available at: http://faostat3.fao.org/home/index.html [Accessed December 21, 2012].

FAO. 2014b. "FAOSTAT." Available at: http://faostat3.fao.org/home/index.html [Accessed September 24, 2014].

Fielding, A.H., and J.F. Bell. 1997. “A review of methods for the assessment of prediction errors in conservation presence/absence models.” *Environmental conservation* 24(01):38–49.

Fischer, Gã¼., M. Shah, F. N. Tubiello, and H. van Velhuizen. 2005. “Socio-economic and climate change impacts on agriculture: an integrated assessment, 1990â€“2080.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 360:2067–2083.

Fischer, G., F.O. Nachtergaele, S. Prieler, E. Teixeira, G. Tóth, H. van Velthuizen, L. Verelst, and D. Wiberg. 2012. “Global Agro-Ecological Zones (GAEZ v3. 0): Model Documentation.” *International Institute for Applied systems Analysis (IIASA), Laxenburg. Rome, Italy: Austria and the Food and Agriculture Organization of the United Nations (FAO).*

Frölicher, T.L., M. Winton, and J.L. Sarmiento. 2013. “Continued global warming after CO2 emissions stoppage.” *Nature Climate Change* 4(1):40–44.

Fujino, J., R. Nair, M. Kainuma, T. Masui, and Y. Matsuoka. 2006. “Multi-gas mitigation analysis on stabilization scenarios using AIM global model.” *The Energy Journal* (Special Issue# 3):343–354.

Gay Garcia, C., F. Estrada, C. Conde, H. Eakin, and L. Villers. 2006. “Potential Impacts of Climate Change on Agriculture: A Case of Study of Coffee Production in Veracruz, Mexico.” *Climatic Change* 79(3-4):259–288.

Ghoshray, A. 2010. “The extent of the world coffee market.” *Bulletin of Economic Research* 62(1):97–107.

Giovannucci, D., B. Lewin, R. Swinkels, and P. Varangis. 2004. “Socialist Republic of Vietnam Coffee Sector Report.” *World Bank, Washington, DC.*

Hannah, L., P.R. Roehrdanz, M. Ikegami, A.V. Shepard, M.R. Shaw, G. Tabor, L. Zhi, P.A. Marquet, and R.J. Hijmans. 2013. “Climate change, wine, and conservation.” *Proceedings of the National Academy of Sciences* 110(17):6907–6912.

Hansen, J., M. Sato, and R. Ruedy. 2012. “PNAS Plus: Perception of climate change.” *Proceedings of the National Academy of Sciences*. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1205276109 [Accessed September 11, 2012].

Harris, R., L.L. Porfirio, S. Hugh, G. Lee, N.L. Bindoff, B. Mackey, and N.J. Beeton. 2013. “To Be Or Not to Be? Variable selection can change the projected fate of a threatened species under future climate.” *Ecological Management & Restoration* 14(3):230–234.

Havlík, P., U.A. Schneider, E. Schmid, H. Böttcher, S. Fritz, R. Skalský, K. Aoki, S.D. Cara, G. Kindermann, F. Kraxner, S. Leduc, I. McCallum, A. Mosnier, T. Sauer, and M. Obersteiner. 2011. “Global land-use implications of first and second generation biofuel targets.” *Energy Policy* 39(10):5690–5702.

Havlik, P., H. Valin, A. Mosnier, M. Obersteiner, J.S. Baker, M. Herrero, M.C. Rufino, and E. Schmid. 2012. “Crop Productivity and the Global Livestock Sector: Implications for

Land Use Change and Greenhouse Gas Emissions." *American Journal of Agricultural Economics* 95(2):442–448.

Hein, L., and F. Gatzweiler. 2006. "The economic value of coffee (Coffea arabica) genetic resources." *Ecological Economics* 60(1):176–185.

Hernandez, P.A., C.H. Graham, L.L. Master, and D.L. Albert. 2006. "The effect of sample size and species characteristics on performance of different species distribution modeling methods." *Ecography* 29(5):773–785.

Heumann, B.W., S.J. Walsh, and P.M. McDaniel. 2011. "Assessing the application of a geographic presence-only model for land suitability mapping." *Ecological Informatics* 6(5):257–269.

Hijmans, R.J. 2012. "Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model." *Ecology* 93(3):679–688.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis. 2005. "Very high resolution interpolated climate surfaces for global land areas." *International journal of climatology* 25(15):1965–1978.

Hylander, K., S. Nemomissa, J. Delrue, and W. Enkosa. 2013. "Effects of Coffee Management on Deforestation Rates and Forest Integrity." *Conservation Biology*:n/a–n/a.

IBGE. 2012. "Sistema IBGE de Recuperação Automática - SIDRA." *IBGE*. Available at: http://sidra.ibge.gov.br/ [Accessed March 28, 2013].

ICO. 2010. "Employment generated by the coffee sector." No. ICC 105-5, ICO. Available at: http://dev.ico.org/documents/icc-105-5e-employment.pdf [Accessed September 29, 2014].

ICO. 2013. "International Coffee Organization - Historical data." *Historical Data*. Available at: http://www.ico.org/new_historical.asp?section=Statistics [Accessed January 29, 2014].

ICO. 2014. "International Coffee Organization - World Coffee Trade." *World Coffee Trade*. Available at: http://www.ico.org/trade_e.asp?section=About_Coffee [Accessed October 2, 2014].

Inman, M. 2011. "Opening the future." *Nature Climate Change* 1(1):7–9.

Instituto Nacional de Estadística y Censos. 2006. "Censo Cafetalero." Available at: http://www.inec.go.cr/Web/Home/GeneradorPagina.aspx [Accessed June 30, 2014].

International Trade Centre. 2011. *The Coffee exporter's guide*.

Jarvis, A., and J. Ramirez. 2010. "Downscaling Global Circulation Model Outputs: The Delta Method." *CIAT Decision and Policy Analysis Working Paper* (1).

Jha, S., C.M. Bacon, S.M. Philpott, V. Ernesto Mendez, P. Läderach, and R.A. Rice. 2014. "Shade Coffee: Update on a Disappearing Refuge for Biodiversity." *BioScience* 64(5):416–428.

Karatzoglou, A., D. Meyer, and K. Hornik. 2006. "Support Vector Machines in R." *Journal of Statistical Software* 15(9). Available at: http://www.jstatsoft.org/v15/i09.

Kuemmerle, T., K. Erb, P. Meyfroidt, D. Müller, P.H. Verburg, S. Estel, H. Haberl, P. Hostert, M.R. Jepsen, T. Kastner, C. Levers, M. Lindner, C. Plutzar, P.J. Verkerk, E.H. van der Zanden, and A. Reenberg. 2013. "Challenges and opportunities in mapping land use intensity globally." *Human settlements and industrial systems* 5(5):484–493.

Kuhn, M. 2008. "Building predictive models in R using the caret package." *Journal of Statistical Software* 28(5):1–26.

Läderach, P., A. Eitzinger, O. Ovalle, S. Carmona, and E. Rahn. 2012. "Future Climate Scenarios for Viet Nam's Robusta Coffee Growing Areas." International Center for Tropical Agriculture.

Läderach, P., A. Eitzinger, O. Ovalle, J. Ramirez, and A. Jarvis. 2010. "Climate Change Adaptation and Mitigation in the Kenyan Coffee Sector." CIAT.

Läderach, P., J. Haggar, C. Lau, A. Eitzinger, O. Ovalle, M. Baca, A. Jarvis, and M. Lundry. 2009. "Mesoamerican coffee: building a climate change adaptation strategy." CIAT Policy brief No. 2, CIAT.

Läderach, P., M. Lundy, A. Jarvis, J. Ramirez, E.P. Portilla, K. Schepp, and A. Eitzinger. 2011. "Predicted impact of climate change on coffee supply chains." In *The Economic, Social and Political Elements of Climate Change*. Springer, pp. 703–723.

Läderach, P., A. Martinez-Valle, G. Schroth, and N. Castro. 2013. "Predicting the future climatic suitability for cocoa farming of the world's leading producer countries, Ghana and Côte d'Ivoire." *Climatic change* 119(3-4):841–854.

Läderach, P., T. Oberthür, S. Cook, M. Estrada Iza, J.A. Pohlan, M. Fisher, and R. Rosales Lechuga. 2011. "Systematic agronomic farm management for improved coffee quality." *Field Crops Research* 120(3):321–329.

Lambin, E.F., and P. Meyfroidt. 2011. "Global land use change, economic globalization, and the looming land scarcity." *Proceedings of the National Academy of Sciences* 108(9):3465–3472.

Lane, A., and A. Jarvis. 2007. "Changes in climate will modify the geography of crop suitability: agricultural biodiversity can help with adaptation." *SAT ejournal* 4(1):1–12.

Lewin, B., D. Giovannucci, and P. Varangis. 2004. "Coffee Markets: New Paradigms in Global Supply and Demand." World Bank. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=996111.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by randomForest." *R news* 2(3):18–22.

Lobo, J.M., A. Jiménez-Valverde, and R. Real. 2008. "AUC: a misleading measure of the performance of predictive distribution models." *Global Ecology and Biogeography* 17(2):145–151.

Merow, C., M.J. Smith, and J.A. Silander. 2013. "A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter." *Ecography*:no–no.

Mitchell, T.D., T.R. Carter, P.D. Jones, M. Hulme, and M. New. 2004. "A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100)." *Tyndall Centre for Climate Change Research Working Paper* 55:25.

Moguel, P., and V.M. Toledo. 1999. "Review: Biodiversity Conservation in Traditional Coffee Systems of Mexico." *Conservation Biology* 13(1):11–21.

Monfreda, C., N. Ramankutty, and J.A. Foley. 2008. "Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000." *Global biogeochemical cycles* 22(1).

Mosnier, A., M. Obersteiner, P. Havlík, E. Schmid, N. Khabarov, M. Westphal, H. Valin, S. Frank, and F. Albrecht. 2014. "Global food markets, trade and the cost of climate change adaptation." *Food security* 6(1):29–44.

Mosnier, A., M. Obersteiner, P. Havlík, H. Valin, and G. Bocqueho. 2012. "Overview of the Globiom model."

Moss, R.H., J. a Edmonds, K. a Hibbard, M.R. Manning, S.K. Rose, D.P. van Vuuren, T.R. Carter, S. Emori, M. Kainuma, T. Kram, G. a Meehl, J.F.B. Mitchell, N. Nakicenovic, K. Riahi, S.J. Smith, R.J. Stouffer, A.M. Thomson, J.P. Weyant, and T.J. Wilbanks. 2010. "The next generation of scenarios for climate change research and assessment." *Nature* 463(7282):747–56.

Müller, C., and R.D. Robertson. 2014. "Projecting future crop productivity for global economic modeling." *Agricultural Economics* 45(1):37–50.

Nakicenovic, N., and R. Swart. 2000. "IPCC Special Report Emission Scenarios."

Nelson, G.C., H. Valin, R.D. Sands, P. Havlík, H. Ahammad, D. Deryng, J. Elliott, S. Fujimori, T. Hasegawa, E. Heyhoe, P. Kyle, M. Von Lampe, H. Lotze-Campen, D. Mason d'Croz, H. van Meijl, D. van der Mensbrugghe, C. Müller, A. Popp, R. Robertson, S. Robinson, E. Schmid, C. Schmitz, A. Tabeau, and D. Willenbockel. 2014. "Climate change effects on agriculture: Economic responses to biophysical shocks." *Proceedings of the National Academy of Sciences* 111(9):3274–3279.

Nix, H.A. 1986. "A biogeographic analysis of Australian elapid snakes." *Atlas of Elapid Snakes of Australia.(Ed.) R. Longmore*:4–15.

Oberthür, T., P. Läderach, H. Posada, M.J. Fisher, L.F. Samper, J. Illera, L. Collet, E. Moreno, R. Alarcón, A. Villegas, H. Usma, C. Perez, and A. Jarvis. 2011. "Regional relationships between inherent coffee quality and growing environment for denomination of origin labels in Nariño and Cauca , Colombia." *Food Policy*.

Van Oijen, M., J. Dauzat, J.-M. Harmand, G. Lawson, and P. Vaast. 2010a. "Coffee agroforestry systems in Central America: I. A review of quantitative information on physiological and ecological processes." *Agroforestry systems* 80(3):341–359.

Van Oijen, M., J. Dauzat, J.-M. Harmand, G. Lawson, and P. Vaast. 2010b. "Coffee agroforestry systems in Central America: II. Development of a simple process-based model and preliminary results." *Agroforestry systems* 80(3):361–378.

Ortega-Huerta, M.A., and A.T. Peterson. 2008. "Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods." *Revista mexicana de biodiversidad* 79(1):205–216.

Otim, S., and P.K. Ngategize. 1993. "Uganda Coffee Supply Response and Export Demand: An Economic Analysis." *African Crop Science Journal* 1(2).

Ovalle, O., P. Läderach, and C. Bunn. 2012. "Global Impact of Climate Change on Coffee Suitability." In *Tropentag*. Tropentag. Göttingen.

Ovalle-Rivera, O., P. Läderach, C. Bunn, M. Obersteiner, and G. Schroth. 2015. "Projected Shifts in Coffea arabica Suitability among Major Global Producing Regions Due to Climate Change" R. D. Loyola, ed. *PLOS ONE* 10(4):e0124155.

Pachauri, R.K., M.R. Allen, V.R. Barros, J. Broome, W. Cramer, R. Christ, J.A. Church, L. Clarke, Q. Dahe, and P. Dasgupta. 2014. "Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change."

Pendergrast, M. 2009. "Coffee second only to oil? Is coffee really the second largest commodity? Mark Pendergrast investigates and finds some startling - Entrepreneur.com." Available at: http://www.entrepreneur.com/tradejournals/article/198849799_1.html [Accessed October 14, 2010].

Pendergrast, M. 2010. *Uncommon grounds: the history of coffee and how it transformed our world*. New York: Basic Books.

Peterson, A.T., M. Pape, and M. Eaton. 2007. "Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent." *Ecography* 30(4):550–560.

Phalan, B., M. Bertzky, S.H.M. Butchart, P.F. Donald, J.P.W. Scharlemann, A.J. Stattersfield, and A. Balmford. 2013. "Crop Expansion and Conservation Priorities in Tropical Countries" S. G. Willis, ed. *PLoS ONE* 8(1):e51759.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. "Maximum entropy modeling of species geographic distributions." *Ecological Modelling* 190(3-4):231–259.

Phillips, S.J., and M. Dudik. 2008. "Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation." *Ecography* 31(2):161–175.

Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. "Newer classification and regression tree techniques: bagging and random forests for ecological prediction." *Ecosystems* 9(2):181–199.

PSAI. 2014. "Philstat." Available at: http://www.philstat.org.ph/ [Accessed May 6, 2014].

Ramirez, J., and A. Jarvis. 2010. "Disaggregation of Global Circulation Model Outputs." *International Center for Tropical Agriculture, CIAT, Cali, Colombia*.

Ramirez-Villegas, J., A.J. Challinor, P.K. Thornton, and A. Jarvis. 2013. "Implications of regional improvement in global climate models for agricultural impact research." *Environmental Research Letters* 8(2):024018.

Ramirez-Villegas, J., A. Jarvis, and P. Läderach. 2013. "Empirical approaches for assessing impacts of climate change on agriculture: the EcoCrop model and a case study with grain sorghum." *Agricultural and Forest Meteorology* 170:67–78.

Randall, D., R. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R. Stouffer, A. Sumi, and K. Taylor. 2007. "Climate Models and Their Evaluation."

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.R-project.org.

Riahi, K., A. Grübler, and N. Nakicenovic. 2007. "Scenarios of long-term socio-economic and environmental development under climate stabilization." *Technological Forecasting and Social Change* 74(7):887–935.

Van Rikxoort, H., G. Schroth, P. Läderach, and B. Rodriguez Sanchez. 2014. "Carbon footprints and carbon stocks reveal climate-friendly coffee production." *Agronomy for Sustainable Development*.

Rosenzweig, C., J. Elliott, D. Deryng, A.C. Ruane, C. Müller, A. Arneth, K.J. Boote, C. Folberth, M. Glotter, N. Khabarov, K. Neumann, F. Piontek, T.A.M. Pugh, E. Schmid, E. Stehfest, H. Yang, and J.W. Jones. 2014. "Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison." *Proceedings of the National Academy of Sciences* 111(9):3268–3273.

Rotter, R.P., T.R. Carter, J.E. Olesen, and J.R. Porter. 2011. "Crop-climate models need an overhaul." *Nature Clim. Change* 1(4):175–177.

Rüegsegger, M. 2008. *Kaffeeanbau in Äthiopien unter verschiedenen Klimaszenarien*. Universität Bern.

Schneider, U.A., P. Havlík, E. Schmid, H. Valin, A. Mosnier, M. Obersteiner, H. Böttcher, R. Skalský, J. Balkovič, T. Sauer, and S. Fritz. 2011. "Impacts of population growth, economic development, and technical change on global food production and consumption." *Agricultural Systems* 104(2):204–215.

Schroth, G., P. Läderach, D.S. Blackburn Cuero, J. Neilson, and C. Bunn. 2014. "Winner or loser of climate change? A modeling study of current and future climatic suitability of Arabica coffee in Indonesia." *Regional Environmental Change*. Available at: http://link.springer.com/10.1007/s10113-014-0713-x [Accessed January 8, 2015].

Schroth, G., P. Läderach, J. Dempewolf, S. Philpott, J. Haggar, H. Eakin, T. Castillejos, J. Garcia Moreno, L. Soto Pinto, R. Hernandez, A. Eitzinger, and J. Ramirez-Villegas. 2009. "Towards a climate change adaptation strategy for coffee communities and ecosystems in the Sierra Madre de Chiapas, Mexico." *Mitigation and Adaptation Strategies for Global Change* 14(7):605–625.

Seale, J.L., A. Regmi, and J. Bernstein. 2003. *International evidence on food consumption patterns*. Economic Research Service, US Department of Agriculture.

Shannon, C.E. 1948. *The mathematical theory of communication*. Citeseer.

Simonett, O. 1999. "Potential Impacts of Global Warming, GRID-Geneva, Case studies on climatic change." *GRID-Geneva Vital Graphics*. Available at: http://www.grida.no/publications/vg/climate/page/3090.aspx [Accessed June 14, 2011].

Simonett, O. 1988. "Using GIS in a Global Context: The Global Resource Information Database (GRID)." In *Proceedings, Eighth Annual ESRI User Conference*.

Skalsky, R., Z. Tarasovicová, J. Balkovic, E. Schmid, M. Fuchs, E. Moltchanova, G. Kindermann, and P. Scholtz. 2008. "GEO-BENE global database for biophysical modeling v. 1.0." *Concepts, methodologies and data*.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller. 2007. "IPCC, 2007: Summary for Policymakers." *Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*.

Soubbotina, T.P. 2004. *Beyond economic growth: an introduction to sustainable development* 2nd ed. Washington, D.C: World Bank.

Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley. 2013. "Climate Change 2013. The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change-Abstract for decision-makers." Groupe d'experts intergouvernemental sur l'evolution du climat/Intergovernmental Panel on Climate Change-IPCC, C/O World Meteorological Organization, 7bis Avenue de la Paix, CP 2300 CH-1211 Geneva 2 (Switzerland).

Tadesse, M., B. Alemu, G. Bekele, T. Tebikew, J. Chamberlin, and T. Benson. 2006. *Atlas of the Ethiopian rural economy*. Intl Food Policy Res Inst.

Thuiller, W. 2004. "Patterns and uncertainties of species' range shifts under climate change." *Global Change Biology* 10(12):2020–2027.

Tropek, R., O. Sedla ek, J. Beck, P. Keil, Z. Musilova, I. imova, and D. Storch. 2014. "Comment on 'High-resolution global maps of 21st-century forest cover change.'" *Science* 344(6187):981–981.

USDA. 2012. "PSD Online - Home." *Production, Supply and Distribution Online*. Available at: http://www.fas.usda.gov/psdonline/ [Accessed December 21, 2012].

VanDerWal, J., L.P. Shoo, C. Graham, and S.E. Williams. 2009. "Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know?" *ecological modelling* 220(4):589–594.

Varangis, P.N. 2003. *Dealing with the coffee crisis in Central America: Impacts and strategies*. World Bank Publications.

Vega, F.E., E. Rosenquist, and W. Collins. 2003. "Global project needed to tackle coffee crisis." *Nature* 425(6956):343.

Vermeulen, S.J., E. Grainger-Jones, and X. Yao. 2014. "Climate change, food security and small-scale producers." *CCAFS Info Brief, CGAIR Research Programme on Climate Change, Agriculture and Food Security (CCAFS), Copenhagen*.

VICOFA. 2004. "Vietnam Coffee-Cocoa Association." Available at: http://www.vicofa.org.vn/.

Van Vuuren, D.P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G.C. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S.J. Smith, and S.K. Rose. 2011. "The representative concentration pathways: an overview." *Climatic Change* 109(1-2):5–31.

Van Vuuren, D.P., M.G. Den Elzen, P.L. Lucas, B. Eickhout, B.J. Strengers, B. van Ruijven, S. Wonink, and R. van Houdt. 2007. "Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs." *Climatic Change* 81(2):119–159.

Ward, R.D. 1911. "The Economic Climatology of the Coffee District of Sao Paulo, Brazil." *Bulletin of the American Geographical Society* 43(6):428–445.

Warszawski, L., K. Frieler, V. Huber, F. Piontek, O. Serdeczny, and J. Schewe. 2014. "The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project framework." *Proceedings of the National Academy of Sciences* 111(9):3228–3232.

Webb, M.S., and V.B. Hall. 2009. "Application Of A Dynamic Panel Data Estimator To Cross-Country Coffee Demand: A Tale Of Two Eras." *Journal of Economic Development* 34(1):1–17.

Weigel, A.P., M.A. Liniger, and C. Appenzeller. 2008. "Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?" *Quarterly Journal of the Royal Meteorological Society* 134(630):241–260.

White, J.W., G. Hoogenboom, B.A. Kimball, and G.W. Wall. 2011. "Methodologies for simulating impacts of climate change on crop production." *Field Crops Research* 124(3):357–368.

Whitsed, R., R. Corner, and S. Cook. 2012. "A model to predict ordinal suitability using sparse and uncertain data." *Applied Geography* 32(2):401–408.

Williams, J.R., and V.P. Singh. 1995. "The EPIC model." *Computer models of watershed hydrology.*:909–1000.

Wintgens, J.N. 2009. *Coffee: Growing, Processing, Sustainable Production - A guidebook for Growers, Processors, Traders, and Researchers* 2nd ed. J. N. Wintgens, ed. Weinheim: Wiley-VCH.

You, L., S. Crespo, Z. Guo, J. Koo, W. Ojo, K. Sebastian, M Tenorio, S. Wood, and U. Wood-Sichra. 2012. "MapSpaM | Home of the Spatial Production Allocation Model." Available at: http://mapspam.info/ [Accessed December 20, 2012].

You, L., and S. Wood. 2006. "An entropy approach to spatial disaggregation of agricultural production." *Agricultural Systems* 90(1):329–347.

You, L., U. Wood-Sichra, S. Fritz, Z. Guo, L. See, and J. Koo. 2014. "Spatial Production Allocation Model (SPAM) 2005 v2.0." Available at: http://mapspam.info [Accessed June 16, 2015].

You, L., S. Wood, and U. Wood-Sichra. 2009. "Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach." *Agricultural Systems* 99(2):126–140.

Zapata-Caldas, E., A. Jarvis, J. Ramirez-Villegas, and C. Lau. 2011. "Potenciales Impactos del Cambio Climático en Cultivos Andinos." Decision and Policy Analysis ProgramA.

Zullo, J., H.S. Pinto, E.D. Assad, and A.M.H. Ávila. 2011. "Potential for growing Arabica coffee in the extreme south of Brazil in a warmer world." *Climatic Change*. Available at: http://www.springerlink.com/index/10.1007/s10584-011-0058-0 [Accessed September 27, 2011].

Zullo Jr, J., H.S. Pinto, and E.D. Assad. 2007. "Impact assessment study of climate change on agricultural zoning." *Meteorological Applications* 13(S1):69.

Zullo Jr, J., H.S. Pinto, E.D. Assad, and S.R. de Medeiros Evangelista. 2008. "Potential Economic Impacts of Global Warming on Two Brazilian Commodities , According to IPCC Prognostics." *Terrae* 3(1):28–39.

Annex 1. Region definitions in Globiom

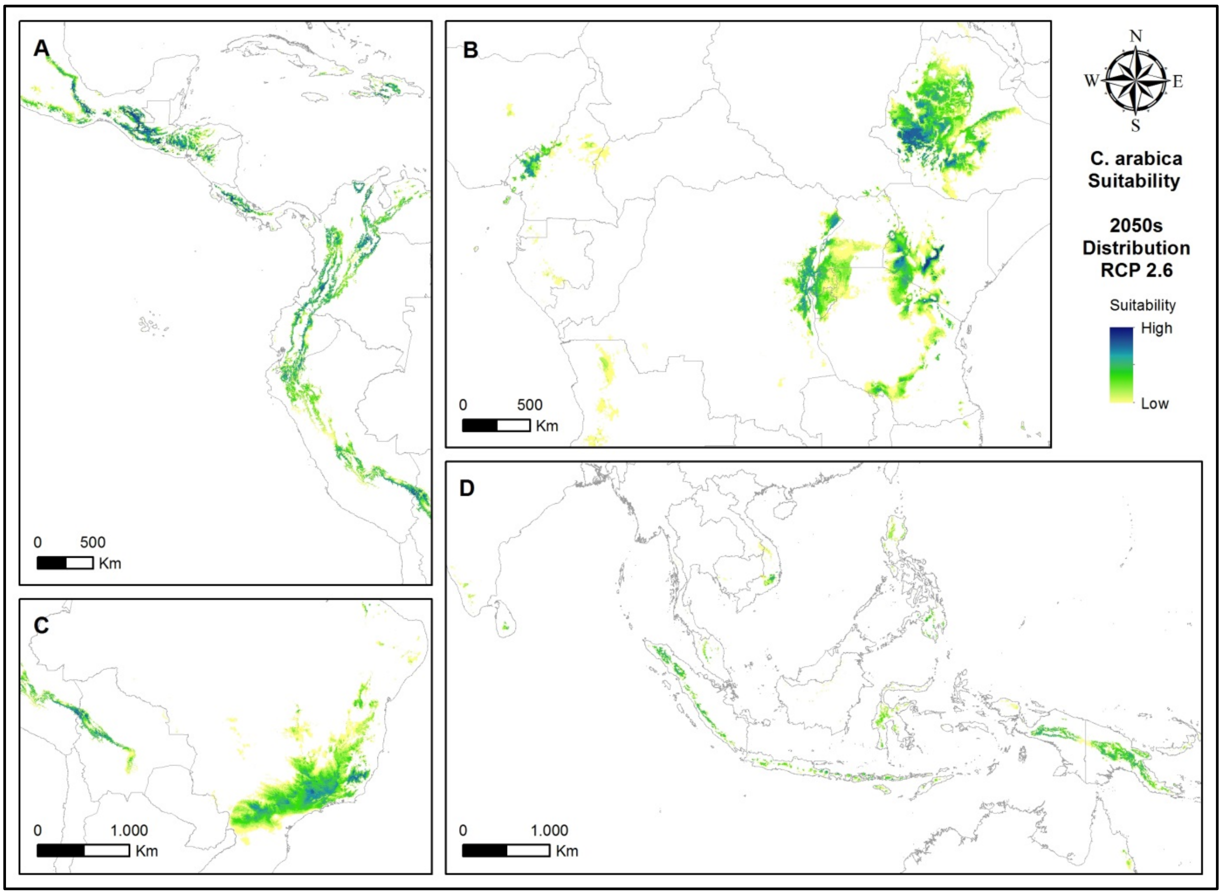
Table S. 1. Region definitions for Globiom and chapter 3

Globiom Model Region	Coffee Region	Countries
CanadaReg		Canada
EU_Baltic		Estonia, Latvia, Lithuania
EU_CentralEast		Bulgaria, CzechRep, Hungary, Poland, Romania, Slovakia, Slovenia
EU_MidWest		Austria, France, Germany, Luxembourg, Netherlands
EU_North		Denmark, Finland, Ireland, Sweden, UK
RCEU		Albania, BosniaHerzg, Croatia, Macedonia, Serbia-Monte
EU_South		Cyprus, Greece, Italy, Malta, Portugal, Spain,
Former_USSR		Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, MoldovaRep, RussianFed, Tajikistan, Turkmenistan, Ukraine, Uzbekistan
JapanReg		Japan
MidEastNorthAfr		Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestin, Qatar, SaudiArabia, Syria, Tunisia, UntdArabEm, WestSahara, Yemen[1]
ROWE		Greenland, Iceland, Norway, Switzerland
TurkeyReg		Turkey
RSEA_OPA	Asian Islands	BruneiDarsm, Indonesia, Malaysia, Myanmar[2], Philippines, Singapore, Thailand[2], TimorLeste
BrazilReg	Brazil	Brazil
RCAM	Central America	Bahamas, Belize, CostaRica, Cuba, DominicanRp, ElSalvador, Guadeloupe, Guatemala, Haiti, Honduras, Jamaica, Nicaragua, Panama, TrinidadTob
CongoBasin	Congo Basin	Cameroon, CentAfrRep, CongoDemR, CongoRep, EqGuinea, Gabon
EasternAf	East Africa	Burundi, Ethiopia, Kenya, Rwanda, Tanzania, Uganda
MexicoReg	North America	Mexico,
USAReg	North America	PuertoRico, USA
ANZ	Pacific Islands	Australia, NewZealand
Pacific_Islands	Pacific Islands	FijiIslands, FrPolynesia, NewCaledonia, PapuaNGuin, Samoa, SolomonIs, Vanuatu
SouthAfrReg	South Africa	SouthAfrica
SouthernAf	South Africa	Angola, Botswana, Comoros, Lesotho, Madagascar, Malawi, Mauritius, Mozambique, Namibia, Reunion, Swaziland, Zambia, Zimbabwe
RSAM	South America	Argentina, Bolivia, Chile, Colombia, Ecuador, FalklandIs, FrGuiana, Guyana, Paraguay, Peru, Suriname, Uruguay, Venezuela
IndiaReg	South Asia	India
RSAS	South Asia	Afghanistan, Bangladesh, Bhutan, Nepal, Pakistan, SriLanka
ChinaReg	South East Asia	China
RSEA_PAC	South East Asia	Cambodia, KoreaDPRp, Laos, Mongolia, VietNam
SouthKorea	South East Asia	KoreaRep
WesternAf	West Africa	Benin, BurkinaFaso, CapeVerde, Chad, CotedIvoire, Djibouti, Eritrea, Gambia, Ghana, Guinea, GuineaBissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, SierraLeone, Somalia, Sudan, Togo

[1]"East Africa" – Coffee region, [2]"South East Asia"-Coffee region (own data and Havlík et al. 2011).

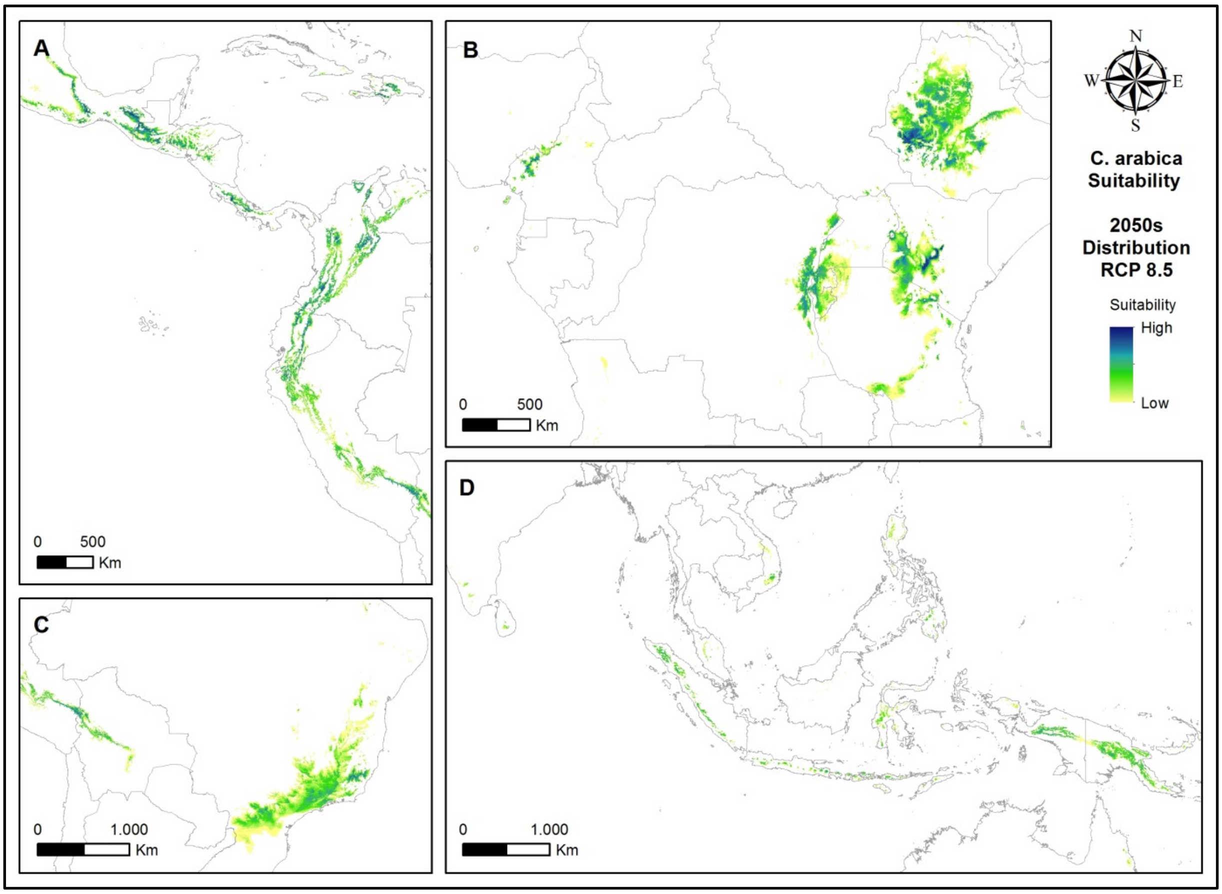
Annex 2. Future coffee suitability

Figure S.1. ***Coffea arabica*** **2050s suitability distribution in the RCP 2.6 scenario**



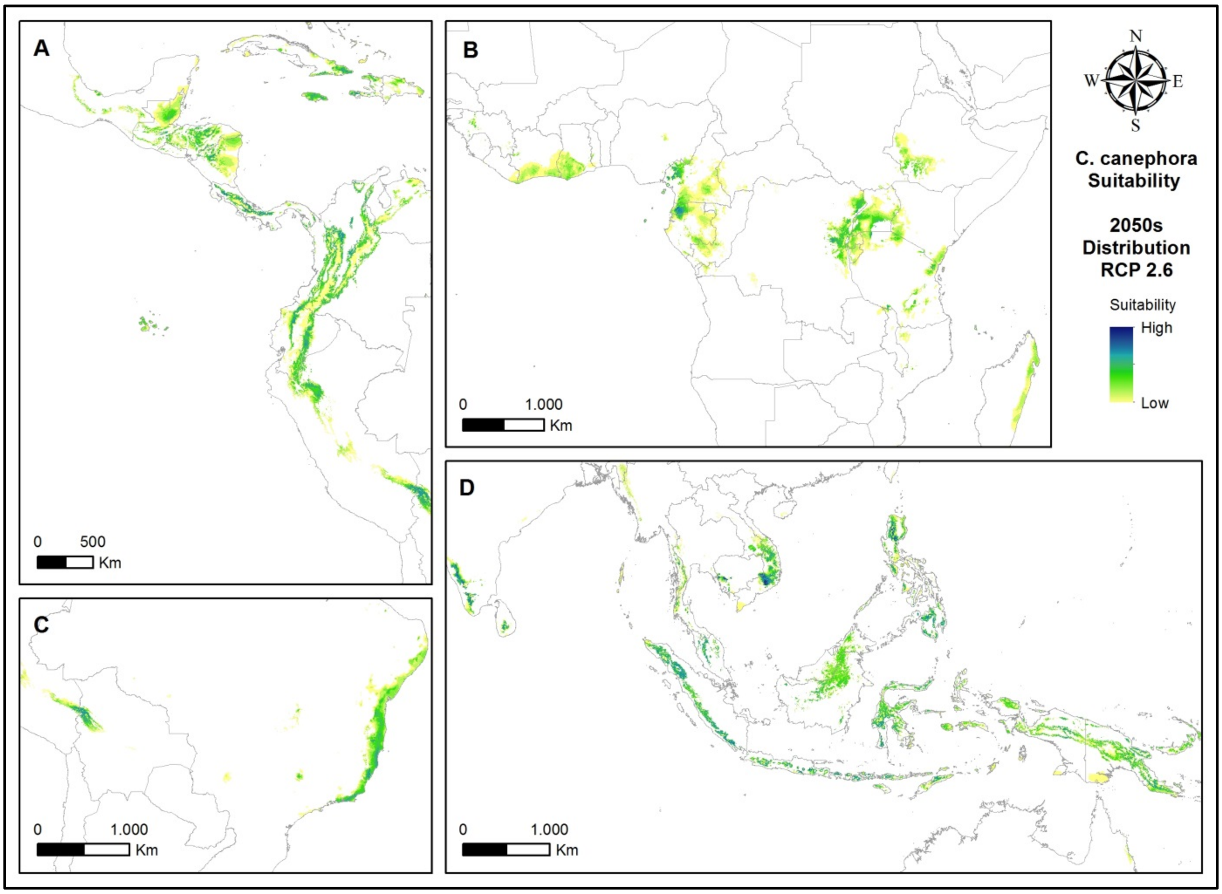
Shown is the mean suitability score of the model ensemble. Dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (Own data and representation).

Figure S.2. ***Coffea arabica*** **2050s suitability distribution in the RCP 8.5 scenario**



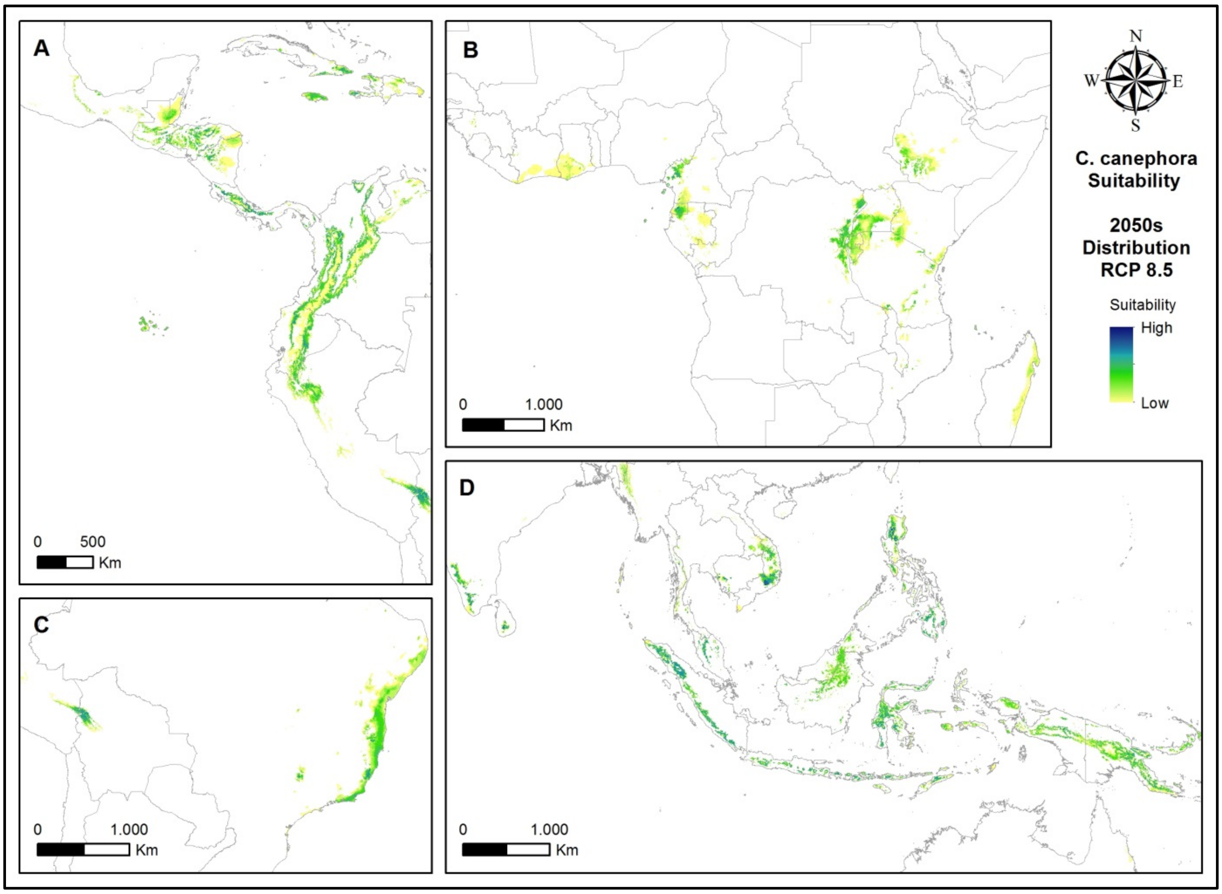
Shown is the mean suitability score of the model ensemble. Dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (Own data and representation).

Figure S.3. ***Coffea canephora*** **2050s suitability distribution in the RCP 2.6 scenario**



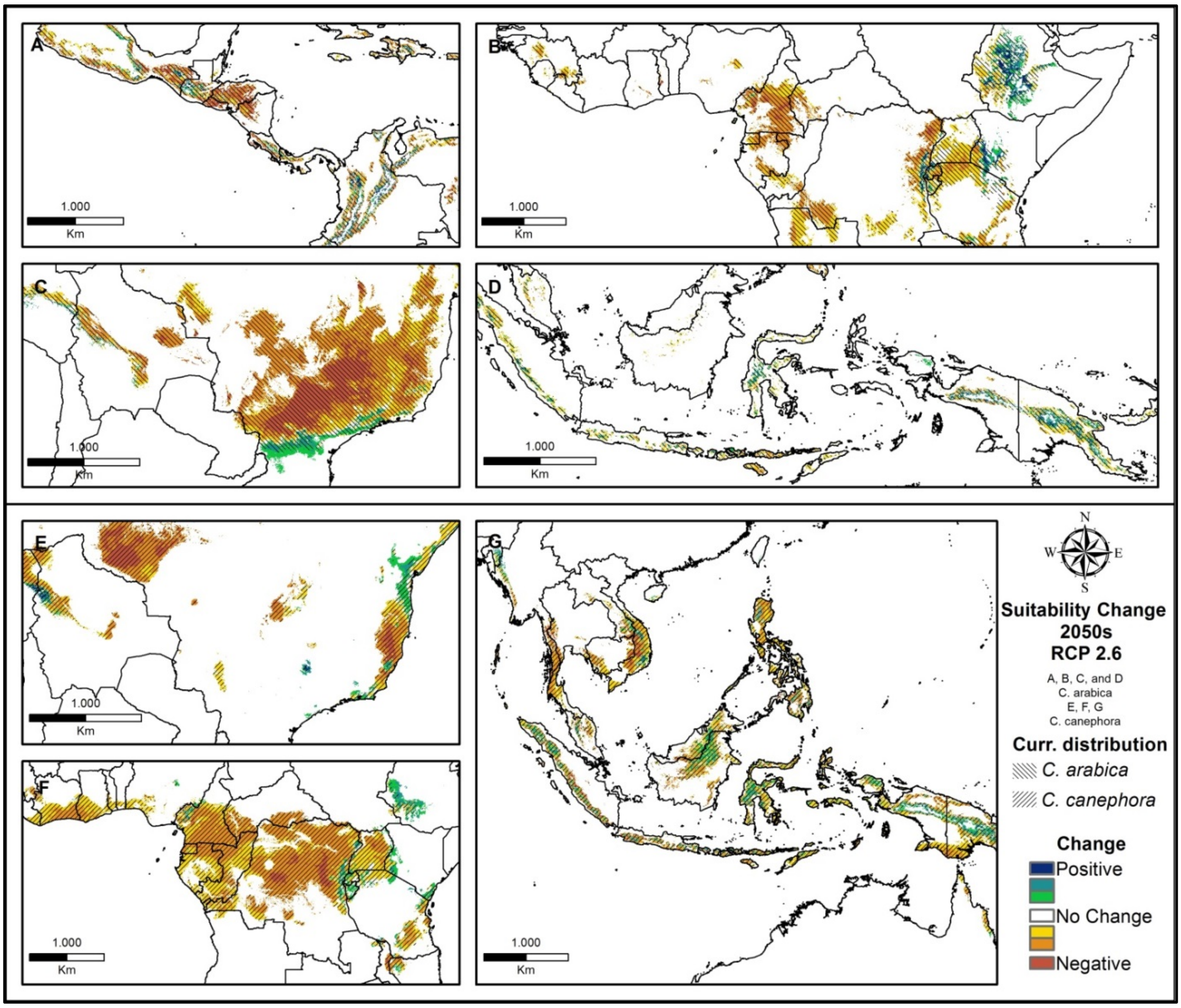
Shown is the mean suitability score of the model ensemble. Dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (Own data and representation).

Figure S.4. ***Coffea canephora*** **2050s suitability distribution in the RCP 8.5 scenario**



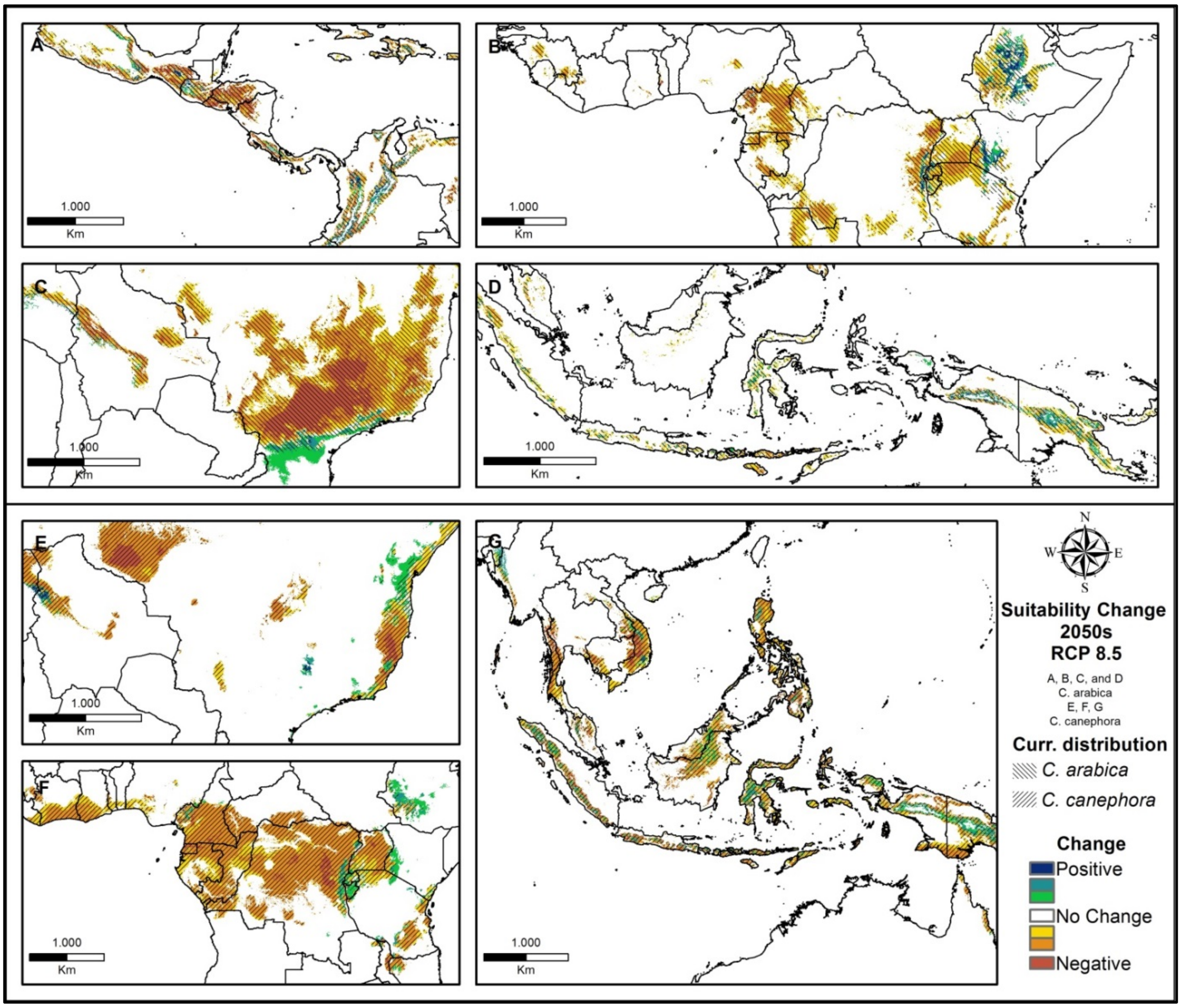
Shown is the mean suitability score of the model ensemble. Dark blue represents highly suitable areas and light yellow marginal suitability. A) Central America and the Andes, B) Central Africa, C) Brazil, D) Asia (Own data and representation).

Figure S.5. Suitability changes by the 2050s in the RCP 2.6 scenario



A-D: Arabica, E-G: Robusta. Hatching indicates the current suitability distribution; Warm colors represent areas with negative climate change impacts and cold colors positive changes

Figure S.6. Suitability changes by the 2050s in the RCP 8.5 scenario



A-D: Arabica, E-G: Robusta. Hatching indicates the current suitability distribution; Warm colors represent areas with negative climate change impacts and cold colors positive changes